CHAPTER 3

Temperature Sensors and Thermal Actuators

The human body and heat

A person consuming 2000 kcal/day, assuming all calories are expended in 24 hours, produces $2000 \times 1000 \times 4.184 = 8.368 \times 10^6$ J of energy. That translates to an average power of $8.368 \times 10^6/24/3600 = 96.85$ W. That's an average power of about 100 W, with lower values during sleep and higher values while active. During exertion such as exercise, a person can produce in excess of 1.5 kW. But the real story is that the body both expends energy and requires the resulting heat to regulate the body temperature and to supply its energy needs. The body requires a fairly narrow range of temperatures. The normal body temperature for most individuals is 37°C, but it can fluctuate somewhat, with women having an average temperature about 0.5°C lower than men. At temperatures above 37°C the body experiences fever, an elevated temperature due to failure of the body to regulate its temperature (typically because of sickness). Hyperthermia is an elevated body temperature that is not a fever, but rather is a reaction to external heat or to drugs or stimulants. Above 41.5°C the body enters into a state called hyperpyrexia, a dangerous state that can lead to death or serious side effects. Lower body temperature, hypothermia, is equally dangerous. Defined as a body core temperature below 35°C, it typically occurs due to exposure to extreme cold or extended immersion in cold water, but can also be due to trauma. Heat regulation in the body is controlled by the hypothalamus in the brain and can be accomplished using a number of methods, including sweating, increasing or decreasing heart rate, shivering, and constriction of blood vessels.

3.1 | INTRODUCTION

Temperature sensors are the oldest sensors in use (excluding the magnetic compass), dating back to the very beginning of the scientific age. Early thermometers were introduced in the early 1600s. Around the middle of the 1600s, the need for standards of temperature measurement were voiced by Robert Boyle and others. Shortly after, about 1700, some temperature scales were already in use, devised by Lorenzo Magalotti, Carlo Renaldini, Isaac Newton, and Daniel Fahrenheit. By 1742, all temperature scales, including the Celsius scale (devised in 1742 by Andres Celsius), but excluding the Kelvin scale, were established. Following the work of Leonard Carnot on engines and heat, Lord Kelvin proposed the absolute scale bearing his name in 1848 and established its relation with the Celsius scale. The temperature scales were further developed and improved until the establishment of the International Practical Temperature Scale in 1927, followed by further revisions to improve accuracy.

The classical thermometer is clearly a sensor, but in its common configuration it does not provide an output signal (but see **Section 3.2**). Its output is read directly. The establishment of temperature sensing was not far behind the development of thermometers. The Seebeck effect was discovered in 1821 by Thomas Johann Seebeck and within a few years it became an accepted method of temperature measurement when Antoine Cesar Becquerel used it in 1826 to build the first modern temperature sensor—the **thermocouple**. The Seebeck effect is the basis of thermocouples and thermopiles, the workhorses of industrial temperature sensing. A second effect related to the Seebeck effect was discovered in 1834 by Charles Athanase Peltier. Named the Peltier effect, it is used for sensing, but more often it is used for cooling or heating as well as for thermoelectric power generation (TEG). Whereas the Seebeck effect has been used from very early on for sensing, applications based on the Peltier effect had to wait for the discovery of semiconductors, in which the effect is much larger than in conductors. These thermoelectric effects are discussed in **Section 3.3** in detail, as well as in **Chapter 4** in conjunction with optical sensing, especially in the infrared region.

Following the observation by Humphry Davey in 1821 that the conductivity of metals decreases with temperature, William Siemens described in 1871 a method for measurement of temperature based on the resistance–temperature relation in platinum. This method has become the basis of **thermoresistive sensors** on which resistance temperature detectors (RTDs) are based. Its extension to semiconductors gave birth to **thermistors** (thermal resistors) as well as to a variety of other useful sensors including semiconductor temperature sensors.

Temperature sensing is both well established and in very widespread use. Many of the sensors available are deceptively simple in construction and can be extremely accurate. On the other hand, they sometimes require special instrumentation and considerable care to achieve this accuracy. A good example is the thermocouple. Arguably the most widespread of temperature sensors, especially at higher temperatures, it is also a very delicate instrument whose output signal is tiny and requires special techniques for measurement, special connectors, elimination of noise, and calibration, as well as a reference temperature. Yet, in its basic construction it is the "mere welding of any two dissimilar conductors" to form a junction. Others seem to be even simpler. A length of copper wire (or any other metal) connected to an ohmmeter can make an instant thermoresistive sensor of surprisingly good quality. Further adding to the widespread use and availability of temperature sensors is the fact that additional physical attributes may be measured indirectly through measurement of temperature. Examples are the use of temperature sensors to measure air velocity or fluid flow (what is measured is the cooling effect of moving air or that of a fluid with reference to a constant temperature) and radiation intensity, especially in the microwave and infrared spectra, by measuring the temperature increase due to absorption of radiation energy.

It should be noted that there are important thermal actuators as well. The fact that metals as well as gases expand with temperature allows these to be used for actuation. In many cases, sensing and actuation can be achieved directly. For example, a column of water, alcohol, or mercury, or a volume of gas will expand to indicate temperature directly or indirectly and, at the same time, this expansion can be used to actuate a dial or a switch. Direct reading cooking thermometers and thermostats are of this type. Another simple example is the bimetal switch commonly used in direction indicators in cars, where the sensing element activates a switch directly without the need of an intermediate controller.

3.1.1 Units of Temperature, Thermal Conductivity, Heat, and Heat Capacity

The SI unit of absolute temperature is the kelvin (K). It is based on the absolute zero. The common unit in day-to-day work is the degree Celsius (°C). The two are the same except for the reference zero temperature. The Kelvin scale's zero is the absolute zero, whereas the Celsius scale is based on the triple point of water. Thus 0°C = 273.15 K and one says that the absolute zero temperature is 0 K or −273.15°C. Conversions between the three common temperature scales are as follows:

From °C to K: N [°C] = ($N + 273.15$) [K]
From °C to °F: P [°C] = ($P \times 1.8 + 32$) [°F]
From K to °C: M [K] = ($M - 273.15$) [°C]
From °F to °C: Q [°F] = ($Q - 32$)/1.8 [°C]
From K to °F: S [K] = ($S - 273.15$) $\times$ 1.8 + 32 [°F]
From °F to K: U [°F] = ($U - 32$)/1.8 + 273.15 [K]

Heat is a form of energy. Therefore its SI derived unit is the joule (J). The joule is a small unit, so units of megajoules (MJ), gigajoules (GJ), and even terajoules (TJ) are used, as are smaller units all the way down to nanojoules (nJ) and even attojoules (aJ). Although there is a whole list of units of energy, some metric and some customary, we mention here only a few of the more common ones. A joule equals a watt-second (1 J = 1 W·s) or a newton meter (N·m), but perhaps in more common use is the kilowatt-hour (kWh) (1 kWh = 3.6 MJ). Another commonly used unit of energy, especially in relation to heat energy is the calorie (cal). The calorie is a thermochemical unit equal to 0.239 J. It should be noted here that in the United States the term calorie usually refers to 1000 calories or the kilocalorie (kcal), that is, what is typically called a calorie in the United States is in fact 1000 cal or 239 J. Neither the calorie nor the watt-hour are SI units. Both are considered obsolete units and their use is discouraged.

Thermal conductivity, denoted as k or λ, is measured in watts per meter per kelvin (W/m/K) and is a measure of the ability of materials to conduct heat. **Heat capacity**, denoted as C, refers to the amount of heat necessary to change the temperature of a substance by a given amount. The SI unit for heat capacity is the joule per Kelvin (J/K). In conjunction with chemical sensors, the **molar heat capacity** is often employed and has units of joules per mole per Kelvin (J/mol/K). Other useful quantities are the **specific heat capacity** (J/kg/K), the amount of heat required to increase the temperature of 1 kg of a substance by 1 K, and **volumetric heat capacity** (J/m^3/K), the amount of heat needed to increase the temperature of 1 m^3 of a substance by 1 K. Often, too, the units of specific heat capacity are given in joules per gram per kelvin (J/g/K) and those of volumetric heat capacity as joules per cubic centimeter per kelvin (J/cm^3/K).

3.2 | THERMORESISTIVE SENSORS: THERMISTORS, RESISTANCE TEMPERATURE SENSORS, AND SILICON RESISTIVE SENSORS

The bulk of thermoresistive sensors may be divided into two basic types: **resistance temperature detectors (RTDs)** and **thermistors** (the name is a concatenation of the words thermal and resistor). RTDs have come to indicate thermoresistive sensors based on solid conductors, usually in the form of metal wires or films. In these devices the resistance of the sensor increases with temperature—that is, the materials used have positive temperature coefficients (PTCs) of resistance. Silicon-based RTDs have also been developed and have the distinct advantage of being much smaller than conductor-based RTDs, as well as exhibiting higher resistances and higher temperature coefficients. Thermistors are semiconductor-based devices and usually have a negative temperature coefficient (NTC) of resistance, but PTC thermistors also exist.

3.2.1 Resistance Temperature Detectors

The earlier sensors of this type were made of an appropriate metal such as platinum, nickel, or copper, depending on the application, temperature range, and often the cost. All RTDs are based on the change in resistance due to the temperature coefficient of resistance (TCR) of the metal being used. The resistance of a conductor of length L with constant cross-sectional area S and conductivity σ (**Figure 3.1**) is

$$R = \frac{L}{\sigma S} \quad [\Omega]. \tag{3.1}$$

The conductivity of the material itself is temperature dependent and given as

$$\sigma = \frac{\sigma_0}{1 + \alpha\,[T - T_0]} \quad [\mathrm{S/m}], \tag{3.2}$$

where α is the temperature coefficient of resistance of the conductor, T is the temperature, and σ_0 is the conductivity of the conductor at the reference temperature T_0. T_0 is usually given at 20°C, but may be given at other temperatures as necessary. The resistance of the conductor as a function of temperature is therefore

$$R(T) = \frac{L}{\sigma_0 S}(1 + \alpha[T - T_0]) \quad [\Omega] \tag{3.3}$$

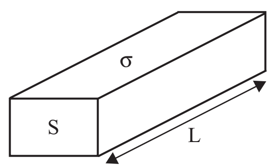


FIGURE 3.1 ■ Geometry used for calculation of resistance of a conductor of length L and uniform cross-sectional area S.

or

$$R(T) = R_0(1 + \alpha[T - T_0]) \quad [\Omega], \tag{3.4}$$

where R_0 is the resistance at the reference temperature T_0. In most cases the temperatures T and T_0 are given in degrees Celsius, but they can be specified in kelvin as well, provided both are given on the same scale.

Although this relation is linear, the coefficient α is usually quite small and the conductivity σ_0 is large. For example, for copper, $\sigma_0 = 5.8 \times 10^7$ S/m and $\alpha = 0.0039$/°C (the coefficient can also be indicated as Ω/Ω/°C) at $T_0 = 20$°C. Taking a wire with a cross-sectional area S of 0.1 mm^2 and length L of 1 m gives a change in resistance of 6.61×10^{-5} Ω/°C and a base resistance of 0.017 Ω at 20°C. This is a change of 0.38%. Thus, for the sensor to be practical, the conductor must be long and thin and/or conductivity must be low. A large temperature coefficient is also useful in that the change in resistance is large and hence processing of the signal obtained is easier. The TCRs and conductivities for a number of useful materials are given in **Table 3.1**.

Equation (3.3) is useful in establishing the physics of RTDs and gives useful insight into how resistance changes with temperature. However, in practice, one typically does not know the properties of the material used in producing the sensor. The typical data available are the nominal resistance of the RTD, usually given at 0°C, the range

TABLE 3.1 ■ Conductivities and temperature coefficients of resistance for selected materials (at 20°C unless otherwise indicated)

Material	Conductivity σ [S/m]	Temperature coefficient of resistance [per °C]
Copper (Cu)	5.8×10^7	0.0039
Carbon (C)	3.0×10^5	−0.0005
Constantan (60% Cu, 40% Ni)	2.0×10^6	0.00001
Chromium (Cr)	5.6×10^6	0.0059
Germanium (Ge)	2.2	−0.05
Gold (Au)	4.1×10^7	0.0034
Iron (Fe)	1.0×10^7	0.0065
Mercury (Hg)	1.0×10^6	0.00089
Nichrome (NiCr)	1.0×10^6	0.0004
Nickel (Ni)	1.15×10^7	0.00672
Platinum (Pl)[2]	9.4×10^6	0.003926 (at 0°C)
Silicon (Si) (pure)	4.35×10^{-6}	−0.07
Silver (Ag)	6.1×10^7	0.0016
Titanium (Ti)	1.8×10^6	0.042
Tungsten (W)	1.8×10^7	0.0056
Zinc (Zn)	1.76×10^7	0.0059
Aluminum (Al)	3.6×10^7	0.0043

Notes:

1. Instead of conductivity, σ [S/m], some sources list resistivity, ρ, measured in ohm-meters ($\rho = 1/\sigma$ [Ωm]).
2. Platinum is a particularly important material and there are different grades with different TCRs in use. The TCR is often given at 0°C. The most common TCRs at 0°C are 0.00385 (European curve), 0.003926 (American curve), and 0.00375 (common in thin-film sensors). The TCR of pure platinum at 0°C is 0.003926. Some alloys in use have TCRs of 0.003916 and 0.003902 (at 0°C). Other grades can be made by alloying pure platinum with materials such as rhodium.
3. The TCR of a material changes with temperature (see **Problem 3.9**). For example, the TCR of pure platinum at 20°C is 0.003729/°C.

EXAMPLE 3.1 Wire-spool sensor

A spool of magnet wire (copper wire insulated with a thin layer of polyurethane) contains 500 m of wire with a diameter of 0.2 mm. It is proposed to use the spool as a temperature sensor to sense the temperature in a freezer. The proposed range is between $-45°C$ and $+10°C$. A milliammeter is used to display the temperature by connecting the sensor directly to a 1.5 V battery and measuring the current through it.

a. Calculate the resistance of the sensor and the corresponding currents at the minimum and maximum temperatures.
b. Calculate the maximum power the sensor dissipates.

Solution: The resistance of a length of wire, disregarding temperature is

$$R = \frac{l}{\sigma s},$$

where l is the length of the wire, S is its cross-sectional area, and σ is its conductivity.

The conductivity is temperature dependent. For copper, **Table 3.1** gives the conductivity as $\sigma_0 = 5.8 \times 10^7$ S/m at 20°C. Thus the resistance is written as a function of temperature using **Equation (3.3)**:

$$R(T) = \frac{l}{\sigma_0 S}(1 + \alpha[T - 20°]).$$

The TCR, α, for copper is given in **Table 3.1**. At $-45°C$,

$$R(-45°) = \frac{500}{5.8 \times 10^7 \times \pi \times (0.0001)^2}(1 + 0.0039[-45 - 20°]) = 204.84\ \Omega.$$

At $+10°C$,

$$R(+10°) = \frac{500}{5.8 \times 10^7 \times \pi \times (0.0001)^2}(1 + 0.0039[10 - 20°]) = 263.7\ \Omega.$$

The resistance changes from 204.84 Ω at $-45°C$ to 263.7 Ω at $+10°C$. The currents are

$$I(-45°) = \frac{1.5}{204.84} = 7.323 \text{ mA}$$

and

$$I(+10°) = \frac{1.5}{263.7} = 5.688 \text{ mA}.$$

The current is linear with temperature and amounts to a sensitivity of 29.72 μA/°C. It is not a particularly large current, but even the simplest digital multimeter should be able to measure a change of 10 μA, or a temperature change of about 0.3°C. With a better microammeter, resolution down to 1 μA, or 0.03°C is possible.

The power dissipated is

$$P(+10^\circ) = I^2R = \left(5.688 \times 10^{-3}\right)^2 \times 263.7 = 8.53\ \text{mW}$$

and

$$P(-45^\circ) = I^2R = \left(7.323 \times 10^{-3}\right)^2 \times 204.84 = 10.98\ \text{mW}.$$

The power is low, an important property in temperature sensors since, as we shall see shortly, the power dissipated in the sensor can lead to errors due to self-heating. Note also the absolute simplicity of this sensor.

EXAMPLE 3.2 Wire RTD resistance and sensitivity

A wire-wound RTD sensor is made of pure platinum wire, 0.01 mm in diameter, to have a resistance of 25 Ω at 0°C. Assume here that the TCR is constant with temperature.

a. Find the necessary length for the wire.
b. Find the resistance of the RTD at 100°C.
c. Find the sensitivity of the sensor in ohms/degree Celsius [Ω/°C].

Solution:
a. The resistance is written as a function of temperature using **Equation (3.3)**:

$$R(T) = \frac{l}{\sigma_0 S}(1 + \alpha[T - 20^\circ]).$$

The TCR, α, for platinum is given in **Table 3.1** at 0°C but the conductivity is given at 20°C. The resistance of the RTD at 0°C is

$$25 = \frac{l}{9.4 \times 10^6 \times \pi \times \left(0.05 \times 10^{-3}\right)^2}(1 + 0.003926[0 - 20]) = 12.48154l \quad [\Omega].$$

This gives

$$l = \frac{25}{12.48154} = 2.003\ \text{m}.$$

The sensor requires 2 m of platinum wire.

b. At 100°C,

$$R(100^\circ\text{C}) = \frac{2.003}{9.4 \times 10^6 \times \pi \times \left(0.05 \times 10^{-3}\right)^2}(1 + 0.003926[100 - 20]) = 35.652\ \Omega.$$

The resistance changes from 25 Ω at 0°C to 35.652 Ω at +100°C.

c. Sensitivity is calculated at an arbitrary temperature by calculating the resistance at that temperature then increasing the temperature by 1°C, calculating the resistance, and subtracting the former from the latter. At a temperature T we have

$$R(T) = \frac{l}{\sigma_0 S}(1 + \alpha[T - 20°]).$$

At a temperature $T + 1$,

$$R(T + 1) = \frac{l}{\sigma_0 S}(1 + \alpha[(T + 1) - 20°]).$$

The difference in resistance between the two is

$$R(T + 1) - R(T) = \frac{l}{\sigma_0 S}(1 + \alpha[T + 1 - 20°]) - \frac{l}{\sigma_0 S}(1 + \alpha[T - 20°]) = \frac{l\alpha}{\sigma_0 S}.$$

Note that this expression is the slope of the function $R(T)$. We get

$$\Delta R = \frac{l\alpha}{\sigma_0 S} = \frac{2.003 \times 0.003926}{9.4 \times 10^6 \times \pi \times \left(0.05 \times 10^{-3}\right)^2} = 0.1065\ \Omega.$$

The sensitivity is therefore 0.1065 Ω/°C.

Check: Since the resistance is linear with temperature, the sensitivity is the same everywhere and thus we can write the resistance at 100°C as

$$R(100°\mathrm{C}) = R(0°\mathrm{C}) + 100 \times \Delta R = 25 + 100 \times 0.1065 = 35.65\ \Omega.$$

The small difference is due to truncation of the numbers during evaluation.

(say between −200°C and +600°C) and additional data on performance such as self-heat, accuracy, and the like. The transfer function can be obtained in one of two ways. First, the manufacturers conform to existing standards that specify the coefficient α that a sensor uses. For example, standard EN 60751, dealing with platinum RTDs, specifies $\alpha = 0.00385$ (this is sometimes called the "European curve"). Other values are 0.003926 ("American curve"), 0.003916, and 0.003902, among others, and relate to grades of platinum (see the notes to **Table 3.1**). This value is part of the sensor's specifications. This allows one to establish an approximate transfer function as follows:

$$R(T) = R(0)[1 + \alpha T] \quad [\Omega], \tag{3.5}$$

where $R(0)$ is the resistance at 0°C and T is the temperature at which the resistance is sought. This is an approximate value because α is itself temperature dependent.

To improve on this, the resistance as a function of temperature is calculated based on relations established from the actual measurements and given, again, by the same standards. This is given as follows:

For $T \geq 0°\mathrm{C}$:

$$R(T) = R(0)[1 + aT + bT^2] \quad [\Omega]. \tag{3.6}$$

The coefficients are calculated for each material based on fixed temperatures (see **Problem 3.6**). For example, for platinum (standard EN 60751, $\alpha = 0.00385$), the coefficients are

$$a = 3.9083 \times 10^{-3},\ \ b = -5.775 \times 10^{-7}.$$

For $T < 0°\text{C}$:

$$R(T) = R(0)[1 + aT + bT^2 + c(T - 100)T^3] \quad [\Omega]. \tag{3.7}$$

The coefficients are (again from standard EN 60751, $\alpha = 0.00385$):

$$a = 3.9083 \times 10^{-3},\ b = -5.775 \times 10^{-7},\ c = -4.183 \times 10^{-12}.$$

These relations are known as the Callendar–Van Dusen equations or polynomials. Instead of using the polynomials one can use design tables that list the values of resistance at various temperatures. For other values of α, the coefficients are different, but are specified by the standards or can be calculated based on accurate measurements. It should be noted that for small sensing spans close to the nominal temperature, the temperature curve is nearly linear and **Equation (3.5)** is sufficiently accurate. **Equations (3.6)** and **(3.7)** are only needed for larger spans or if sensing is done at low or high temperatures (see **Example 3.3**).

EXAMPLE 3.3 **RTD representation and accuracy**

An RTD with nominal resistance of 100 Ω at 0°C is specified for the range −200°C to +600°C. The engineer has the option of using the approximate transfer function in **Equation (3.5)** or the exact transfer function in **Equations (3.6)** and **(3.7)**. Assume $\alpha = 0.00385/°\text{C}$.

a. Calculate the error incurred by using the approximate transfer function at the extremes of the range.
b. What are the errors if the range used is −50°C to +100°C?

Solution:
a. From **Equation (3.5)**:

At 600°C,

$$R(600°\text{C}) = R(0)[1 + \alpha T] = 100\,[1 + 0.00385 \times 600] = 331\ \Omega.$$

At −200°C,

$$R(-200°\text{C}) = 100[1 - 0.00385 \times 200] = 23\ \Omega.$$

From **Equation (3.6)**:

$$\begin{aligned} R(600°\text{C}) &= R(0)[1 + aT + bT^2] \\ &= 100\,[1 + 3.9083 \times 10^{-3} \times 600 - 5.775 \times 10^{-7} \times 600^2] = 313.708\ \Omega. \end{aligned}$$

From **Equation (3.7)**:

$$\begin{aligned} &100[1 + 3.9083 \times 10^{-3} \times (-200) - 5.775 \times 10^{-7} \times 200^2 \\ &\quad - 4.183 \times 10^{-12} \times (-200)^3] = 18.52\ \Omega. \end{aligned}$$

The resistance calculated with the approximate formula is higher by 5.51% at 600°C and higher by 24.19% at −200°C. These deviations are not acceptable and therefore one cannot use the approximate formula for the whole range—the use of the Callendar–Van Dusen relations is essential.

b. From **Equation (3.5)**:

At 100°C,

$$R(100°\mathrm{C}) = R(0)[1 + \alpha T] = 100\,[1 + 0.00385 \times 100] = 138.5\ \Omega.$$

At −50°C,

$$R(-50°\mathrm{C}) = 100[1 - 0.00385 \times 50] = 80.75\ \Omega.$$

From **Equation (3.6)**:

$$R(100°\mathrm{C}) = 100\,[1 + 3.9083 \times 10^{-3} \times 100 - 5.775 \times 10^{-7} \times 100^2] = 138.5055\ \Omega.$$

From **Equation (3.7)**:

$$\begin{aligned}&100[1 + 3.9083 \times 10^{-3} \times (-50) - 5.775 \times 10^{-7} \times 50^2\\&\quad - 4.183 \times 10^{-12} \times (-50)^3] = 80.3063\ \Omega.\end{aligned}$$

The resistance calculated with the approximate formula is only 0.0397% lower at 100°C and lower by 0.552% at −50°C. These deviations may be quite acceptable and the approximate formula can be used.

In designing RTDs, one has to be careful to minimize the effect of tension or strain on the wires. The reason for this is that tensioning a conductor changes its length and cross-sectional area (constant volume), which has exactly the same effect on resistance as a change in temperature. An increase in strain on the conductor increases the resistance of the conductor. This effect will be explored in detail in **Chapter 6** when we talk about strain gauges. There we shall see that the opposite problem occurs—changes in temperature cause errors in strain readings and these errors must be compensated for.

A characteristic property of wire RTDs is their relatively low resistance. High resistances would require very long wires or excessively thin wires. Another consideration is cost. High resistance RTDs require more material, and since most RTDs are based on platinum, material costs can be significant. For wire sensors, a satisfactory resistance is from a few ohms to a few tens of ohms, but thin film sensors with higher resistances can be made. In wire sensors, the wire is made of a fairly thin uniform wire wound in a small diameter coil (typically, but not always) and the coil is then supported on a suitable support such as mica or glass. If the total length of the wire is small, this support is not necessary and the wire coil, or sometimes just a length of wire, may be free standing or threaded around pegs to keep it in place. Depending on the intended use, the wire and support may be enclosed in an evacuated glass (typically Pyrex) tube with connecting wires going through the tube or in a highly conductive metal (sometime stainless steel) to allow better heat transfer to the sensing wire and therefore faster response of the sensor, or in a ceramic enclosure for higher temperature applications. They may also be flat for surface applications, encircling, and other forms.

When precision sensors are needed, platinum or platinum alloy is the first choice because of its excellent mechanical and thermal properties. In particular, platinum is chemically stable even at elevated temperatures, it resists oxidation, can be made into thin wires of high chemical purity, resists corrosion, and can withstand severe environmental conditions. For these reasons it can be used at temperatures up to about 850°C and down to below −250°C. On the other hand, platinum is very sensitive to strain and to chemical contaminants, and because its conductivity is high, the wire length needed is long (a few meters, depending on the required resistance). As a consequence, the resulting sensor is physically large and not suitable for sensing where temperature gradients are high.

For less demanding applications, both in terms of stability and temperature, nickel, copper, and other conducting materials offer less expensive alternatives at reduced performance. Nickel can be used from about −100°C to about 500°C, but its R-T curve is not as linear as that of other materials. Copper has excellent linearity, but can only be used from about −100°C to about 300°C at best. At higher temperatures, tungsten is often a good choice.

Thin film sensors are produced by depositing a thin layer of a suitable material, such as platinum or one of its alloys, on a thermally stable, electrically nonconducting ceramic that must also be a good heat conductor. The thin film can then be etched to form a long strip (typically in a meander fashion) and the sensor is potted in epoxy or glass to protect it. The final package is typically small (a few millimeters), but can vary in size depending on the application and required resistance. Typical resistance is 100 Ω, but much higher resistances, upwards of 2000 Ω are available. Thin film sensors are typically small and relatively inexpensive and are often the choice in modern sensors, especially when the very high precision of platinum wire sensors is not needed. **Figure 3.2** shows schematic constructions of wire and thin film RTDs.

RTDs typically come in relatively low resistances, especially in the case of wire RTDs. Because of this, an important issue in precision sensing is the resistance of the lead wires, which, necessarily, are made of other materials compatible with external circuits (copper, tinned copper, etc.). The resistance of the lead wires also changes with temperature, and these effects can add to errors in the sensing circuit since these resistances are not negligible (except in some high-resistance thin-film RTDs). Because of this, some commercial sensors come in two-, three-, or four-wire configurations as shown in **Figure 3.3**. The purpose of these configurations is to facilitate compensation for the lead wires. We shall see how this is done in **Chapter 11**, but it should be noted here that the two-wire sensor's leads cannot be compensated, whereas the three- and four-wire sensors allow compensation of the lead resistance and should be used when high precision is essential. To understand why these configuration are important, it is

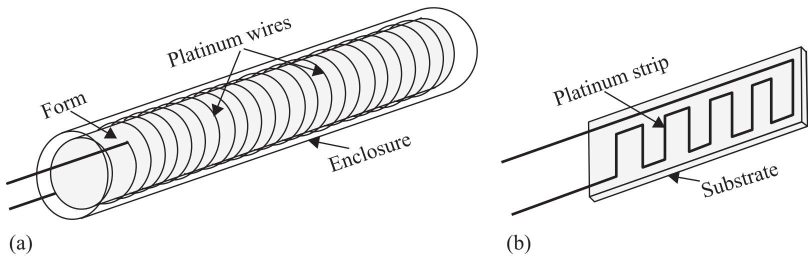


FIGURE 3.2 ■ Schematic construction of RTDs. (a) Wire-wound RTD. (b) Thin film RTDs.

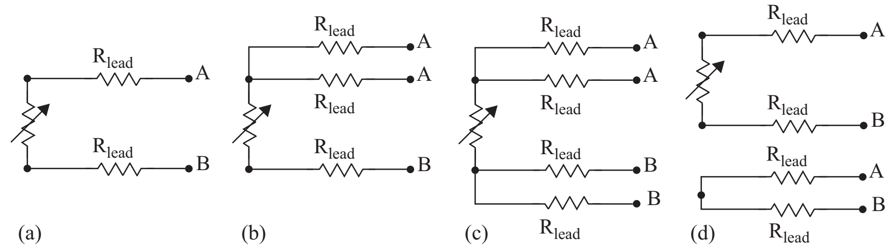


FIGURE 3.3 ■ RTD connection styles. (a) Two-wire (uncompensated). (b) Three-wire. (c) Four-wire. (d) Two-wire with compensation loop. The three- and four-wire styles allow compensation for temperature variations and resistance of the lead wires.

sufficient to note that if the resistance in *A-A* in **Figure 3.3b** is subtracted from the total resistance (measured between *A-B*), one obtains the resistance of the RTD regardless of the resistance of the lead wires. Similar compensation can be achieved with the configuration in **Figure 3.3c** and **3.3d**, but we will see in **Chapter 11** that other methods are often more effective and easier to implement.

Thermoresistive sensors must be calibrated for operation in the range of temperatures for which they were designed. Calibration procedures and calibration temperatures are specified in standards.

The accuracy of thermoresistive sensors can vary considerably depending on materials, temperature range, construction, and methods of measurement. Typical accuracies are on the order of ±0.01°C to ±0.05°C. Higher- and lower-accuracy sensors are available.

The stability of RTDs is measured in degrees Celsius per year (°C/year) and is on the order of 0.05°C/year or less for platinum sensors. Other materials have poorer stability.

3.2.1.1 Self-Heat of RTDs

We shall discuss the connection of sensors in measuring circuits in **Chapter 11**. At this point, however, it is important to mention the fact that many temperature sensors, including thermoresistive sensors, are very much subject to errors due to increases in their own temperature produced by the heat generated in them by the current used to measure their resistance. This is of course a problem with any active sensor, but it is particularly acute in temperature sensors. The rise in temperature may be understood qualitatively from the fact that the higher the current in the sensor, the larger the output signal available. This is particularly important for wire sensors whose resistances are small. On the other hand, power dissipated in the conductor is proportional to the square of the current, and this power can raise the temperature of the sensor, introducing an error. The power can be calculated quantitatively as $P_d = I^2R$, where I is the current (DC or RMS) and R is the resistance of the sensor. In many sensors the power dissipated follows a much more complicated relation so that the relation between current and temperature increase can be quite complex. Typically, as part of the specification of the sensor, the temperature increase per unit power (°C/mW in most cases) is given by the manufacturer, allowing the designer to compensate for these errors in the reading of the sensor. Typical errors are on the order of 0.01°C/mW to 0.2°C/mW, depending on the sensor and on environmental factors such as the cooling conditions (moving air or standing air, contact with heat sinks, in stationary or moving fluids, etc.).

EXAMPLE 3.4 Self-heat of RTDs

Consider the self-heat of an RTD operating in the range −200°C to +850°C that has a nominal resistance of 100 Ω at 0°C and a temperature coefficient of resistance of 0.00385/°C. Its self-heat is provided in its data sheet as 0.08°C/mW in air (typically this value is given at a low airspeed of 1 m/s). Calculate the maximum error expected due to self-heat if

a. The resistance is measured by applying a constant voltage of 0.1 V across the sensor.
b. The resistance is measured by applying a constant current of 1 mA through the sensor.

Note: Both of these measurements provide a nominal current of 1 mA at 0°C.

Solution: First we need to calculate the resistances at the extremes of the span using **Equations (3.6)** and **(3.7)**. These provide the following values:

$$\begin{aligned}R(-200^\circ\mathrm{C}) &= R(0)[1 + aT + bT^2 + c(T-100)T^3]\\ &= 100[1 + 3.9083\times10^{-3}\times(-200) - 5.775\times10^{-7}\times 200^2\\ &\quad -4.183\times10^{-12}\times(-200)^3]\\ &= 18.52\ \Omega\end{aligned}$$

and

$$\begin{aligned}R(850^\circ\mathrm{C}) &= R(0)[1 + aT + bT^2]\\ &= 100[1 + 3.9083\times10^{-3}\times 850 - 5.775\times10^{-7}\times 850^2]\\ &= 390.48\ \Omega.\end{aligned}$$

a. For a constant voltage source, we write the power dissipated as follows:

At −200°C and 850°C,

$$P(-200^\circ\mathrm{C}) = \frac{V^2}{R} = \frac{0.01}{18.52} = 0.54\ \mathrm{mW}$$

and

$$P(850^\circ\mathrm{C}) = \frac{V^2}{R} = \frac{0.01}{390.48} = 0.0256\ \mathrm{mW}.$$

The maximum error occurs at −200°C and equals

$$\text{error at } -200^\circ\mathrm{C} = \frac{0.54}{0.08} = 6.75^\circ\mathrm{C}$$

and

$$\text{error at } 850^\circ\mathrm{C} = \frac{0.0256}{0.08} = 0.32^\circ\mathrm{C}.$$

At the high end of the span the error is only 0.32°C.

b. With a current source, we write

$$P(-200°\text{C}) = I^2R = \left(1 \times 10^{-3}\right)^2 \times 18.52 = 0.0185 \text{ mW}$$

and

$$P(850°\text{C}) = I^2R = \left(1 \times 10^{-3}\right)^2 \times 390.48 = 0.39 \text{ mW}.$$

The maximum error occurs at 850°C and equals

$$\text{error at } 850°\text{C} = \frac{0.39}{0.08} = 4.875°\text{C}.$$

At the low end of the range the error is only 0.23°C:

$$\text{error at } -200°\text{C} = \frac{0.0185}{0.08} = 0.23°\text{C}.$$

Both methods are used and in both the errors vary with temperature. To reduce the error the current can be reduced, but it cannot be too small or difficulties in measurements as well as noise may be encountered.

3.2.1.2 Response Time

The response of most temperature sensors is slow, especially if they are physically large. Typically these are on the order of a few seconds (90% of steady state) and are given in the specification data published by manufacturers. It can range from as little as 0.1 s in water to 100 s in air. It also changes with flowing or standing water or air and these data are usually available. Wire RTDs have the slowest response because of their physical size. Typical specifications are for 50% and 90% of steady-state response in moving air and flowing water (see **Example 3.5**). The response time may be specified for other response levels and under other conditions as necessary. The response time is measured by applying a step change in temperature, ΔT, and measuring the time it takes the sensor to reach a certain temperature (usually one measures the sensor's resistance to deduce its temperature). For example, 50% of steady state means the sensor has reached a temperature equal to its initial value before the step change has been applied plus 50% of the step. The time needed to reach this temperature is the 50% response time.

EXAMPLE 3.5 **Specification of response time in RTDs**

The response time for the sensor in **Example 3.4** is evaluated experimentally in moving air and in flowing water as follows: For the measurement in air, the RTD is placed in a stream of air at an ambient temperature of 24°C moving at approximately 1 m/s. At a time $t = 0$, a heater is turned on, heating the air to 50°C. For the flowing water test, the RTD is placed in a pipe and allowed to settle at the ambient temperature of 24°C. A stream of cold water at 15°C moving at approximately 0.4 m/s is turned on at $t = 0$. The data obtained are shown in the tables below. The resistance of the sensor was measured and its temperature calculated from these data.

RTD in moving air

Time [s]	0	1	2	3	4	5	6	7	8	9	10
Temperature [°C]	24	25	26.4	28.6	31.6	35	38.3	40.5	42.1	43.5	44.4
Time [s]	11	12	13	14	15	16	17	18	19	20	
Temperature [°C]	45.6	46	46.6	47.1	47.5	47.7	48	48.2	48.5	48.8	

RTD in moving water

Time [s]	0	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5
Temperature [°C]	24	23.1	21.7	20.3	18.8	17.6	16.8	15.9	15.6	15.3	15.2
Time [s]	.55	.6	.65	.7	.75	.8	.85	.9	.95	1.0	
Temperature [°C]	15.1	15	15	15	15	15	15	15	15	15	

a. Estimate the 50% and 90% response time in air.

b. Estimate the 50% and 90% response time in water. A 50% and 90% response means that the sensor has reached 50% or 90% of the final reading expected.

Solution: In many cases the data will be given in plots, but in this case we have tabulated data allowing direct calculation of the response time.

a. 50% of steady state for the measurement in air means $24 + (50 - 24) \times 0.5 = 37°C$. 90% of steady state means $24 + (50 - 24) \times 0.9 = 47.4°C$.

The response time may be estimated from the tables above (using linear interpolation between tabulated values):

- The 50% response time is approximately 5.5 s.
- The 90% response time is approximately 14.75 s.

b. In water, the step is negative, equal to $-9°C$. Therefore 50% of steady state means $24 + (15 - 24) \times 0.5 = 19.5°C$ and 90% of steady state means $24 + (15 - 24) \times 0.9 = 15.9°C$. Using the table above and interpolating between values, we get

- the 50% response time is approximately 0.23 s.
- the 90% response time is approximately 0.35 s.

3.2.2 Silicon Resistive Sensors

The conductivity of semiconductors is best explained in terms of quantum effects. We will discuss the quantum effects in semiconductors in the next chapter in conjunction with the photoconductive effect, but for the purpose of this discussion it is useful to point out some of the thermal effects that affect conductivity. To do so we use the classical model of valence and conduction electrons and known relations in semiconductors. Valence electrons may be viewed as those bound to atoms and hence not free to move. Conduction electrons are free to move and affect the current through the semiconductor. In a pure semiconductor, most electrons are valence electrons and are said to be in the valence band. For an electron to move into the conduction band it must acquire additional energy, and in moving into the conduction band it leaves behind a hole (positively charged particle). This energy is called the band gap energy and is material dependent. In the case we are dealing with here, the additional energy comes from heat, but of course it can come from radiation (light, nuclear, electromagnetic). Based on this description, the higher the temperature, the higher the number of electrons (and holes) available and hence the higher the current that can flow through the device (i.e., the lower its resistance). As temperature increases, the resistance decreases and hence pure semiconductors such as silicon typically have negative temperature coefficient (NTC) characteristics. Of course, the behavior of the material is much more complex than that given here and must take into account many parameters such as carrier mobility (which in itself may be temperature dependent) and the purity level of the semiconductor.

Semiconductors are rarely used as pure (intrinsic) materials. More often, impurities are introduced into the intrinsic material in a process called doping. Specifically, when doping silicon with an *n*-type impurity such as arsenic or antimony, the reverse effect is observed below a certain temperature. Typically for *n*-type silicon, a positive temperature coefficient (PTC) is observed below about 200°C. Above that temperature the properties of intrinsic (pure) silicon predominate and the behavior is NTC. The explanation is that at these higher temperatures the energy is so high as to spontaneously generate carriers (move them into the conduction band) regardless of doping. Of course, for practical sensing applications using silicon semiconductor devices, the range of interest is below 200°C.

The conductivity of semiconductors is given by the following relation:

$$\sigma = e(n\mu_e + p\mu_h) \quad [\mathrm{S/m}], \tag{3.8}$$

where e is the charge of the electron (1.61×10^{-19} C), n and p are the concentrations of the electrons and holes in the material (in units of particles/cm^3), and μ_e and μ_p are the mobilities of the electrons and holes, respectively (typically given in units of cm^2/V/s). This relation clearly indicates that conductivity depends on the type of material, since both concentrations and mobilities are material dependent. They are also temperature dependent. In an intrinsic material, the concentration of electrons and holes is equal ($n = p$), but of course in doped materials they are not. However, the concentrations are related through the **mass-action law**:

$$np = n_i^2, \tag{3.9}$$

where n_i is the intrinsic concentration. As the concentration of one type of carrier increases through doping, the other decreases proportionally. In the limit, one type of carrier dominates and the material becomes an *n*-type or *p*-type material. In that case the conductivity of the semiconductor is

$$\sigma = en_d\mu_d \quad [\mathrm{S/m}], \tag{3.10}$$

where n_d is the concentration of the dopant and μ_d is its mobility.

The relations for conductivity apply to all semiconductors and can be used to calculate conductivity based on the properties of the semiconductor. In particular, the concentrations of carriers are temperature dependent and that relation is nonlinear. Nevertheless, the conductivity variation with temperature is a useful measure of temperature. A class of temperature sensors based on silicon in which the nonlinearity is

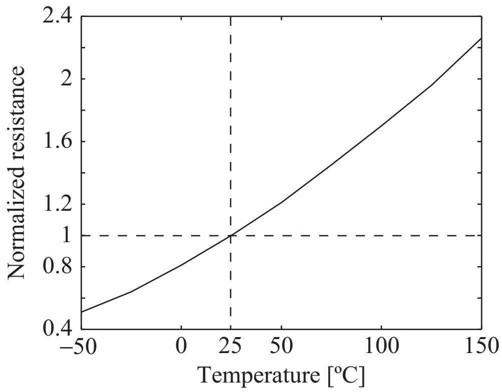


FIGURE 3.4 ■ Normalized resistance (at 25°C) versus temperature for a silicon resistive sensor. The nominal resistance of 1 kΩ at 25°C is indicated by the intersection of the dashed lines.

relatively mild exists. These are called **silicon resistive sensors**. The reduction in nonlinearity is based on construction and the proper selection of dopants, allowing for a useful sensor with sensitivity much higher than that of metal-based sensors, but, as expected, with a much narrower range of temperatures.

Silicon resistive sensors are somewhat nonlinear and offer sensitivities on the order of 0.5−0.7%/°C. They can operate in a limited range of temperatures like most semiconductor devices based on silicon (between −55°C to +150°C). Physically these sensors are very small, made of a piece of silicon with two electrodes deposited on it and encapsulated, usually in epoxy or glass. Typical resistances are on the order of 1 kΩ, specified at a temperature in the span of the device (typically 25°C). Because of the problem of self-heat, the current through these sensors must be kept to a minimum. As a whole, these devices are simple and inexpensive, but their accuracy is limited, with most devices exhibiting errors between 1% and 3%. The normalized resistance of a silicon resistive sensor is shown in **Figure 3.4**.

The transfer function of silicon resistive sensors is given by the manufacturer as a table, as a plot (similar to the one shown in **Figure 3.4**), or as a polynomial. These specification are for individual sensors or for families of sensors and depend on construction and materials. The resistance may be written in general using the Callendar–Van Dusen equation:

$$R(T) = R(0)[1 + a(T - T_{\text{ref}}) + b(T - T_{\text{ref}})^2 + HOT] \quad [\Omega], \tag{3.11}$$

where *HOT* (higher-order term) is a correction term that may be added to accommodate specific curves, especially at high temperatures. The coefficients are derived from the response of the specific sensor being evaluated (see **Example 3.6**).

EXAMPLE 3.6 Silicon resistive sensor

A silicon resistive sensor is described by the first two terms in **Equation (3.11)** with coefficient $a = 7.635 \times 10^{-3}$, $b = 1.731 \times 10^{-5}$ with a reference resistance of 1 kΩ at 25°C. The sensor is to be used for temperature sensing in the range 0°C–75°C. Calculate the maximum deviation of resistance from linearity, where the linear response is given by **Equation (3.4)** with a temperature coefficient of 0.013/°C.

Solution: It is best to calculate and plot the response to see the behavior. The response is plotted in **Figure 3.5**. Clearly the maximum error in output (resistance) is at the lowest temperature.

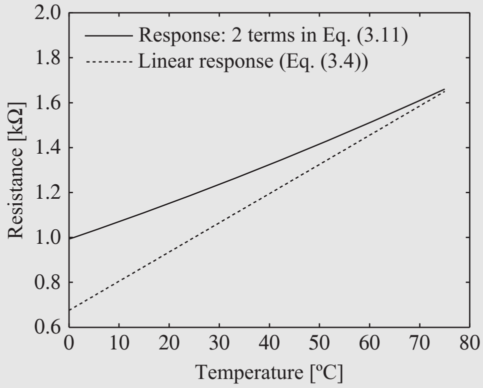


FIGURE 3.5 ■ Response of a silicon resistive sensor showing the deviation between the approximate linear formula and the second-order formula.

At 0°C using **Equation (3.4)**:

$$R(0°\text{C}) = R_0(1 + \alpha[T - T_0]) = 1000(1 + 0.013[0 - 25]) = 675.0\ \Omega.$$

Using **Equation (3.11)**:

$$\begin{aligned} R(0°\text{C}) &= R_0[1 + a(T - T_{\text{ref}}) + b(T - T_{\text{ref}})^2] \\ &= 1000\left(1 + 7.635 \times 10^{-3}[0 - 25] + 1.731 \times 10^{-5}[0 - 25]^2\right) \\ &= 992.4\ \Omega. \end{aligned}$$

The difference is 317.4 Ω. This is 31.98% and clearly the linear formula cannot be used for practical implementation. In application of a sensor of this type, the resistance of the sensor is measured and the temperature is calculated from **Equation (3.11)**, perhaps through the use of a lookup table stored in a microprocessor or direct evaluation of the expressions.

3.2.3 Thermistors

Thermistors (**therm**al res**istors**) came into existence together with other semiconductor devices and have been used for temperature sensing starting in the 1960s. As their name implies, these are thermal resistors, made of semiconducting metal oxides that have high temperature coefficients. Most metal oxide semiconductors have NTCs and their resistance at reference temperatures (typically 25°C) can be rather high. A simple model of a thermistor is the following:

$$R(T) = R_0 e^{\beta(1/T - 1/T_0)} = R_0 e^{-\beta/T_0} e^{\beta/T} \quad [\Omega], \tag{3.12}$$

where R_0 [Ω] is the resistance of the thermistor at the reference temperature T_0, β [K] is the **material constant** and is specific for the particular material used in a device, $R(T)$ is the resistance of the thermistor and T is the temperature sensed [K]. This relation is clearly nonlinear and is only approximate. The inverse relation to **Equation (3.12)** is also useful, particularly in evaluating temperature from measured resistance, and is often used in sensing:

$$T = \frac{\beta}{\ln(R(T)/R_0 e^{-\beta/T_0})} \quad [\mathrm{K}]. \tag{3.13}$$

The model in **Equation (3.12)** can be improved by using the **Steinhart–Hart** equation, which gives the resistance as

$$R(T) = e^{\left(x-\frac{y}{2}\right)^{1/3} - \left(x+\frac{y}{2}\right)^{1/3}} \quad [\text{in } \Omega], \quad y = \frac{a - 1/T}{c}, \quad x = \sqrt{\left(\frac{b}{3c}\right)^3 + \frac{y^2}{4}}. \tag{3.14}$$

The constants a, b, and c are evaluated from three known points on the thermistor response. The inverse relation is often used as well:

$$T = \frac{1}{a + b\ln(R) + c\ln^3(R)} \quad [\mathrm{K}]. \tag{3.15}$$

Equations (3.12) and **(3.14)** establish approximate transfer functions for the thermistor suitable for many applications.

Because the variation between devices can be large, it is often necessary to establish the transfer function through calibration, as discussed in **Chapter 2**. In most cases **Equation (3.15)** is used to evaluate the coefficients a, b, and c and only then does one use **Equation (3.14)** to write the resistance as a function of temperature. The coefficients are often evaluated by manufacturers of thermistors and are available in tables for use in calibration. Similarly, if the simplified form of the transfer function is used, one starts with **Equation (3.13)**, from which the term β may be easily evaluated.

Thermistors are available in higher accuracy models through the processes of trimming the device. The method of production itself can also affect the transfer function of the device. Thermistors are produced by a number of methods. Bead thermistors (**Figure 3.6a**) are essentially a small volume of the metal oxide with two conductors (platinum alloy for high-quality thermistors, copper or copper alloy for inexpensive devices) attached through a heat process. The bead is then coated, usually with glass or epoxy. Another method is to produce chips with surface electrodes (**Figure 3.6b**) that are then connected to leads and the device encapsulated. Chips can be easily trimmed for specific resistance values. A third method of production is to deposit the semiconductor on a substrate following standard semiconductor production methods (see **Figure 3.7b**). These are particularly useful in integrated devices and in complex sensors such as radiation sensors. While the method of production is important, perhaps more important is the encapsulation, since that is the main means of ensuring long-term stability. This encapsulation is one of the main differences between various thermistors.

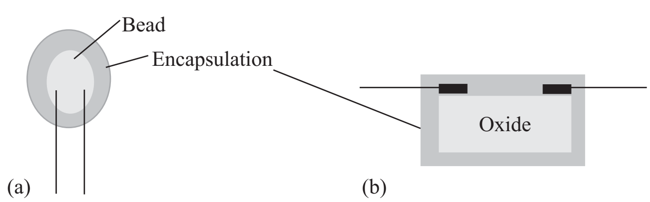


FIGURE 3.6 ■ (a) Construction of a bead thermistor. (b) Construction of a chip thermistor.

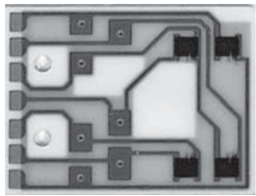
(a)  (b) 

FIGURE 3.7 ■ (a) Two common types of thermistors. (b) Deposited thermistors on a ceramic substrate (the thermistors are the four dark, rectangular areas on the right).

Good thermistors are glass encapsulated, while epoxy potting is common for less expensive devices. Sometimes stainless steel jackets are added for protection in harsh environments. **Figure 3.7a** shows a picture of two epoxy-encapsulated bead thermistors and two chip thermistors. The size of the device is also dictated by the production method (with bead thermistors being the smallest) and this dictates the thermal response of the device. Typically the thermal response of thermistors is relatively short, primarily because of their small physical size.

Although most thermistors are NTC devices, PTC devices can also be made from special materials. These are usually based on barium titanate ($BaTiO_3$) or strontium titanate ($SrTiO_3$) with the addition of doping agents that make them semiconducting. These materials have very high resistances and a highly nonlinear transfer function. Nevertheless, in a small useful range they exhibit a mildly nonlinear curve with a PTC of resistance. Unlike NTC thermistors, PTC thermistors have a steep curve in the useful range (large changes in resistance) and are therefore more sensitive than NTCs in that range. Overall, PTC thermistors are not as common as NTC thermistors, but they have one advantage that is common to all PTC devices (including wire-wound sensors). If connected to a voltage source, as the temperature increases, the current decreases, and therefore they cannot overheat due to self-heating. This is an intrinsic protection mechanism that can be very useful in high-temperature applications. In contrast, NTC thermistors can overheat under the same conditions.

Thermistors exhibit errors due to self-heating similar to those of RTDs. Typical values are between 0.01°C/mW in water and 1°C/mW in air. However, since thermistors are available in a wide range of resistances that can reach a few megaohms with high sensitivity, currents through thermistors are typically very low and self-heating is not usually a problem. On the other hand, some thermistors can be very small, increasing the effect. There are instances in which a thermistor is deliberately heated by passing a current through it, taking advantage of its self-heating properties. An example of this will be discussed in **Chapter 6**. Self-heating is specified by the manufacturer in a manner analogous to that given for RTDs.

The long-term stability of thermistors has been an issue in the past since all thermistors exhibit changes in resistance with aging, especially immediately after production. For this reason they are aged prior to shipment by maintaining them at an elevated temperature for a specified period of time. Good thermistors exhibit negligible drift after the aging process, allowing accurate measurements on the order of 0.25°C with excellent repeatability.

The temperature range of thermistors is higher than that of silicon RTDs and can exceed 1500°C and reach down to about −270°C. The thermistor is often the sensor of

EXAMPLE 3.7 NTC thermistor

A thermistor has a nominal resistance of 10 kΩ at 25°C. To evaluate the thermistor, the resistance at 0°C is measured as 29.49 kΩ. Calculate and plot the resistance of the thermistor between −50°C and +50°C.

Solution: Using **Equation (3.12)** we must first evaluate the coefficient β. This is done as follows:

$$R(0°\mathrm{C}) = 29{,}490 = 10{,}000e^{\beta(1/(273.15)-1/(273.15+25)}\quad [\Omega].$$

To calculate β, we write

$$\ln\left(\frac{29{,}490}{10{,}000}\right) = \beta(1/(273.15) - 1/(298.15))$$

or

$$\beta = \ln\left(\frac{29{,}490}{10{,}000}\right)\left(\frac{1}{(1/(273.15) - 1/(298.15))}\right) = 3.523 \times 10^3\ \mathrm{K}.$$

Now the expression for resistance is

$$R(T) = 10{,}000e^{3.523\times10^3\times(1/T-1/(298.15)}\quad [\Omega].$$

The resistance at −50°C is 530,580 Ω and at + 50°C is 4008 Ω:

$$R(T = 223.15\ \mathrm{K}) = 10{,}000e^{3.523\times10^3\times(1/223.15-1/(298.15)} = 530{,}580\ \Omega$$

$$R(T = 323.15\ \mathrm{K}) = 10{,}000e^{3.523\times10^3\times(1/323.15-1/(298.15)} = 4008\ \Omega.$$

Note the nonlinear behavior of the thermistor over this wide range (**Figure 3.8**). Also, because this is an approximation, the narrower the span, the better the approximation, provided that β is calculated (or given) at a temperature within the span. Clearly this also assumes that β is independent of temperature.

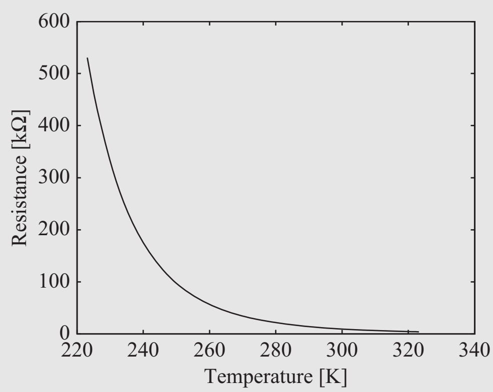


FIGURE 3.8 ■ Response of a thermistor between −50°C (223.15 K) and 50°C (323.15 K).

choice in many consumer products because of the low cost, small size, and simple interfacing needed to make it work.

3.3 | THERMOELECTRIC SENSORS

As indicated in the introduction, thermoelectric sensors are among the oldest sensors, some of the most useful and most commonly used, and have been in use for well over 150 years. And yet, at first sight, this seems curious since the signals produced by thermoelectric sensors are small and difficult to measure and are plagued by noise problems. Perhaps the main reason for their success, particularly in the early years, is the fact that these are passive sensors—they generate electric emfs (voltages) directly and hence all one needs to do is measure the voltage. In the early years, in the absence of amplifiers and controllers, one could still measure the emf, though small, and get an accurate reading of the temperature. Also, they can be produced by anyone with minimum skill. They have other properties that have ensured transcendence into the modern era. In addition to being well developed, simple, rugged, and inexpensive, thermoelectric sensors can operate on almost the entire practically useful range of temperatures from near absolute zero to about 2700°C. No other sensor technology (other than perhaps infrared temperature sensors) can match even a fraction of this range.

There is really only one type of thermoelectric sensor, often called a **thermocouple**. However, there are variations in nomenclature and in construction. Thermocouple usually refers to a junction made of two dissimilar conductors. A number of these junctions connected in series is referred to as a **thermopile**. Semiconductor thermocouples and thermopiles have similar functions but also serve in the reverse function, to generate heat or to cool, and can therefore be used as actuators. These devices are usually called **thermoelectric generators** (TEGs), or sometimes **Peltier cells,** indicating their actuation use, but they can be used as sensors.

Thermocouples are based on the Seebeck effect, which in turn is the sum of two other effects—the Peltier effect and the Thomson effect. These two effects and the resulting Seebeck effect can be described as follows:

The **Peltier effect** is heat generated or absorbed at the junction of two dissimilar materials when an emf exists across the junction due to current in the junction. The effect occurs either by connecting an external emf across the junction or it may be generated by the junction itself, depending on the mode of operation. In either case a current must flow through the junction. This effect has found applications in cooling and heating, particularly in portable refrigerators and in cooling electronic components. Discovered in 1834 by Charles Athanase Peltier, it was developed into its current state in the 1960s as part of the space program. The devices in existence have benefited considerably from developments in semiconductors, and particularly high-temperature semiconducting materials.

The **Thomson effect**, discovered in 1892 by William Thomson (Lord Kelvin), functions such that a current-carrying wire, if unevenly heated along its length, will either absorb or radiate heat depending on the direction of current in the wire (from hot to cold or from cold to hot).

The **Seebeck effect** is an emf produced across the junction between two dissimilar conducting materials. If both ends of the two conductors are connected and a temperature difference is maintained between the two junctions, a thermoelectric current will flow through the closed circuit (**Figure 3.9a**). Alternatively, if the circuit is opened (**Figure 3.9b**), an emf will appear across the open circuit. It is this exact emf that is measured in a thermocouple sensor. The effect was discovered in 1821 by Thomas Johann Seebeck.

In the following simplified analysis we assume that the two junctions in **Figure 3.9b** are at different temperatures, T_1 and T_2, and the conductors are homogeneous. We can then define the Seebeck emf across each of the conductors, *a* and *b*, as

$$emf_a = \alpha_a(T_2 - T_1) \quad \text{and} \quad emf_b = \alpha_b(T_2 - T_1). \tag{3.16}$$

In these relations α_A and α_B are the absolute Seebeck coefficients given in microvolts per degree Celsius (μV/°C) and are properties of the materials involved. The thermoelectric emf generated by a thermocouple made of two wires, A and B, is therefore

$$emf_T = emf_a - emf_b = (\alpha_a - \alpha_b)(T_2 - T_1) = \alpha_{ab}(T_2 - T_1). \tag{3.17}$$

The term α_{ab} is the relative Seebeck coefficient of the material combination *a* and *b* (**Table 3.3**). These coefficients are available for various material combinations and indicate the sensitivity of the thermocouple. Some are listed in **Table 3.3**. Other relative Seebeck coefficients may be obtained from the absolute coefficients in **Table 3.2** and similar tables for other materials by subtracting the absolute coefficients one from the other.

The Seebeck coefficients are rather small—on the order of a few microvolts per degree to a few millivolts per degree for the largest coefficients. This means that in many cases the output from the thermocouple will have to be amplified before it can be

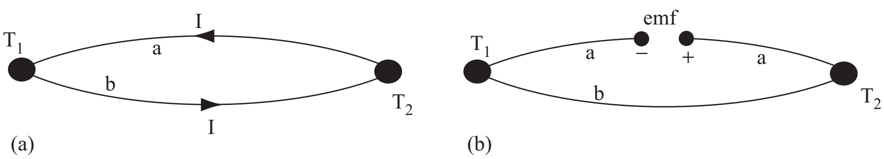


FIGURE 3.9 ■ (a) A thermoelectric current flows in a circuit comprised of two junctions at different temperatures. (b) An emf is developed across the open circuit.

TABLE 3.2 ■ Absolute Seebeck coefficients for selected elements (thermoelectric series)

Material	α [μV/K]
p-Silicon	100–1000
Antimony (Sb)	32
Iron (Fe)	13.4
Gold (Au)	0.1
Copper (Cu)	0
Silver (Ag)	−0.2
Aluminum (Al)	−3.2
Platinum (Pt)	−5.9
Cobalt (Co)	−20.1
Nickel (Ni)	−20.4
Bismuth (Sb)	−72.8
n-Silicon	−100 to −1000

TABLE 3.3 ■ Relative Seebeck coefficients for some material combinations

Materials	Relative Seebeck coefficient at 25°C [μV/°C]	Relative Seebeck coefficient at 0°C [μV/°C]
Copper/constantan	40.9	38.7
Iron/constantan	51.7	50.4
Chromel/alumel	40.6	39.4
Chromel/constantan	60.9	58.7
Platinum (10%)/rhodium-platinum	6.0	7.3
Platinum (13%)/rhodium-platinum	6.0	5.3
Silver/palladium	10	
Constantan/tungsten	42.1	
Silicon/aluminum	446	
Carbon/silicon carbide	170	

used in practical applications. This also implies that special care must be taken in connecting to thermocouples to avoid noisy signals and errors due to, for example, induced emfs from to external sources. Direct measurement of the output, which was the main method of using these sensors in the past, is still possible if used strictly for temperature measurement and no further processing of the output is needed. More often, however, the signal will be used to take some action (turn on or off a furnace, detect a pilot flame before turning on the gas, etc.), and that implies at least some signal conditioning and a controller to affect the action.

The operation of thermocouples is based on three laws—the thermoelectric laws—that summarize the discussion above. These are

1. **Law of homogeneous circuit:** ***A thermoelectric current cannot be established in a homogeneous circuit by heat alone***. This law establishes the need for junctions of dissimilar materials since a single conductor is not sufficient to establish an emf and hence a current.
2. **Law of intermediate materials:** ***The algebraic sum of the thermoelectric forces (emfs) in a circuit composed of any number and combination of dissimilar materials is zero if all junctions are at the same temperature***. This establishes the fact that additional materials may be connected in the thermoelectric circuit without affecting the output of the circuit as long as any junctions added to the circuit are kept at the same temperature. Also, the law indicates that voltages are additive so multiple junctions may be connected in series to increase the output (thermopiles).
3. **Law of intermediate temperatures:** ***If two junctions at temperatures T_1 and T_2 produce Seebeck voltage V_2 and temperatures T_2 and T_3 produce Seebeck voltage V_1, then temperatures T_1 and T_3 produce Seebeck voltage $V_3 = V_1 + V_2$***. This law establishes methods of calibration for thermocouples.

Note: Some sources list five laws, looking at more detailed behavior. The three laws listed here are inclusive and describe all observed effects.

Based on the principles described above, thermocouples are usually used in pairs (but there are exceptions and variations) so that one junction is at the sensing temperature while the second is at a reference temperature, usually a lower temperature, but it can also be a higher temperature. This is shown in **Figure 3.10**, where the voltmeter

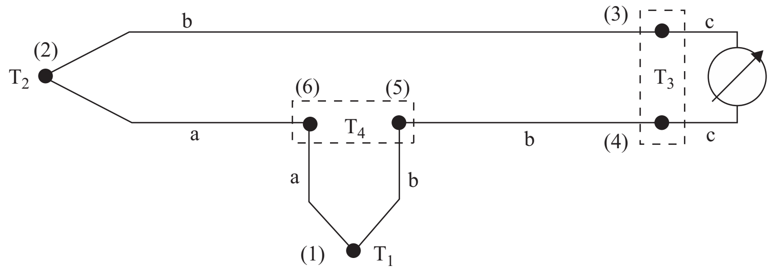


FIGURE 3.10 ■ A measuring thermocouple (hot junction 2) and reference thermocouple (cold junction 1) and additional junctions introduced by connections.

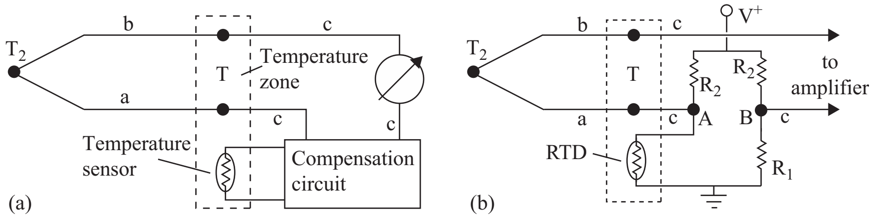


FIGURE 3.11 ■ (a) Connection of a thermocouple through a uniform temperature zone (T) and a compensation circuit to replace the reference junction. (b) The compensation circuit.

represents the device to which the sensor is connected (usually an amplifier). Any connection in the circuit between dissimilar materials adds an emf due to that junction. However, any pair of junctions at identical temperatures may be added without changing the output. In **Figure 3.10** the output is produced by junctions (1) and (2) for the following reason: junctions (3) and (4) are identical (one between material b and c and one between material c and b) and their temperature is the same. Thus no net emf due to this pair is produced. Junctions (5) and (6) also produce zero emf since they are junctions between identical materials that produce zero emf at any temperature. Note that each connection (to the reference junction and to the measuring instrument) necessarily adds two junctions. This then indicates the strategy in sensing: any junction that is not sensed or is not a reference junction must either be between identical materials or must come in pairs and both junctions in the pair must be maintained at the same temperature. In addition, it is a good precaution to use unbroken wires leading from the sensor to the reference junction or to the measuring instrument. If splicing is necessary to extend the length, identical wires must be used to avoid additional emfs.

Connection of thermocouples can be done in many ways, each with their own advantages. One of the most common connections is shown in **Figure 3.11a**. The two junctions (between material b and c and between material a and c) are placed in a so-called **uniform temperature zone** or **isothermal zone**. This can be a small junction box or may in fact be just two junctions in very close proximity—as long as they can be guaranteed to be at identical temperatures. In this case there is no cold junction, but a compensation circuit is added to ensure that the potentials on the junction b–c and a–c, together with the compensation potential, "simulate" the effect of the cold junction. The compensation circuit guarantees that the output is zero at the reference temperature (usually 0°C).

If a reference junction as in **Figure 3.10** is used, it is critical that the temperature is known and constant. In such cases the temperature of the reference junction may be measured separately by a sensor (usually an RTD, sometimes a thermistor, but in any case not a thermocouple) and the reading used to compensate for any change in the temperature of the reference junction. The reference junction may be held at the temperature of a water–ice mixture that guarantees a temperature of 0°C (some variations from this can occur if the water is contaminated or if the atmospheric pressure changes). The temperature of the ice bath should be monitored even if no specific compensation is incorporated in the circuits. An alternative to the ice–water mixture is boiling water, with the same precautions being taken. These two temperatures are also commonly used for calibration of thermocouples. In normal operation, use of an ice–water mixture or boiling water is inconvenient at best. In many applications, the method in **Figure 3.11a** is used, which does not require a reference junction or a fixed temperature and hence avoids the errors involved in the use of a reference junction, not to mention the difficulty of maintaining a known constant temperature. It does, however, require measurement of the temperature in the temperature zone and a compensation circuit that supplies the equivalent emf expected from the reference junction. Further, the temperature sensor cannot be a thermocouple. The measured emf in **Figure 3.11a** is as follows:

$$emf = \alpha_{\mathrm{ba}} T_2 - [\alpha_{\mathrm{bc}} + \alpha_{\mathrm{ca}}]T + emf_{\mathrm{comp}} \quad [\mu\mathrm{V}], \tag{3.18}$$

where α_{ba}, α_{bc}, and α_{ca} are the relative Seebeck coefficients. Of the terms in **Equation (3.18)**, $\alpha_{\mathrm{ba}} T_2$ is of interest and the term $[\alpha_{\mathrm{bc}} + \alpha_{\mathrm{ca}}]T$ may be viewed as part of the reference emf, measured at the temperature *T*. This term may be written in terms of the absolute Seebeck coefficients of the three materials as follows:

$$[\alpha_{bc} + \alpha_{ca}]T = [(\alpha_b - \alpha_b) + (\alpha_c - \alpha_a)]T = (\alpha_b - \alpha_a)T = \alpha_{ba} T. \tag{3.19}$$

Equation (3.18) now becomes

$$emf = \alpha_{\mathrm{ba}} T_2 - \alpha_{\mathrm{ba}} T + emf_{\mathrm{comp}} \quad [\mu\mathrm{V}]. \tag{3.20}$$

The compensation term emf_{comp} is added to ensure that T_2 is correctly measured. The purpose of the compensation circuit is to cancel the term $\alpha_{\mathrm{ba}} T$. That is, the emf supplied by the compensation circuit must be

$$emf_{\mathrm{comp}} = \alpha_{ba} T \quad [\mathrm{mV}], \tag{3.21}$$

where *T* is the temperature of the temperature zone. Under these conditions the measured emf is $emf = \alpha_{\mathrm{ba}} T_2$ and is entirely due to T_2. It should also be noted that the coefficient α_{ba} is the relative Seebeck coefficient of the sensing junction, so that compensation is based on the sensitivity of the sensing junction alone.

To understand the method of compensation, consider **Figure 3.11b**. The cold junction is replaced with a potential difference V_{BA}. The resistance R_1 is selected to equal the resistance of the RTD at 0°C, denoted as R_T. The resistors R_2 are equal and are selected to produce a potential difference per degree Celsius as required by the type of thermocouple being used. Typically the RTD is a platinum RTD with a resistance around 100 Ω and the reference potential V^+ is regulated at an arbitrary level, but typically between 5 V and 12 V. The potential at point A is

$$V_A = \frac{V^+}{R_1 + R_2} R_1. \tag{3.22}$$

The potential at point B is temperature dependent as follows:

$$V_{\mathrm{B}} = \frac{V^{+}}{R_2 + R_T(1+\alpha T)} R_T(1+\alpha T), \tag{3.23}$$

where α is the resistance temperature coefficient of the RTD in use and R_T is its resistance at 0°C. The potential that replaces the cold junction is

$$emf_{\mathrm{comp}} = V_{\mathrm{BA}} = \frac{V^{+}}{R_2 + R_T(1+\alpha T)} R_T(1+\alpha T) - \frac{V^{+}}{R_1 + R_2} R_1. \tag{3.24}$$

This relation allows immediate calculation of the resistance R_2 since, for any given type of thermocouple, V_{BA} is known for any temperature T. **Example 3.8a** shows how the actual calculation of resistance is done.

The method of compensation discussed above does have a slight shortcoming: the range of temperatures of the temperature zone cannot be too far from the reference temperature of the RTD. The reason for this is that the compensation circuit is designed to produce zero emf at the reference temperature of the RTD (as can be verified from the equations above and from **Figure 3.11b**). As long as the temperature range is relatively small, the method is very accurate (see **Example 3.8**).

EXAMPLE 3.8 Cold junction compensation of a K-type thermocouple

Consider the cold junction compensation of a chromel-alumel thermocouple using a platinum RTD, as shown in **Figure 3.11b**. The RTD has a resistance of 100 Ω at 0°C and has a TCR coefficient of 0.00385. The relative Seebeck coefficient (sensitivity) for the K-type thermocouple at 0°C is 39.4 μV/°C (see **Table 3.3**).

a. Given a regulated voltage source of 10 V, calculate the resistance R_2 required for this type of thermocouple.
b. Calculate the error in temperature measurement at 45°C if the temperature zone is at $T = 27$°C and explain the source of this error.

Solution: We use **Equation (3.24)** directly. However, since the Seebeck coefficient is given per degree Celsius, the temperature T can be taken as any value above (or below) 0°C. We will take it as 1°C for convenience. With $R_1 = 100\ \Omega$, we have

a.

$$39.4 \times 10^{-6} = \frac{10}{R_2 + 100(1 + 0.00385 \times 1)} 100(1 + 0.00385 \times 1) - \frac{10}{100 + R_2} 100$$

or

$$39.4 \times 10^{-6} = \frac{1003.85}{R_2 + 100.385} - \frac{1000}{100 + R_2}.$$

Cross-multiplying and separating R_2 we get

$$R_2^{\,2} - 97515.351 R_2 + 10038.5 = 0.$$

Solving this equation gives

$$R_2 = 97{,}515.3\ \Omega.$$

We will take the value, 97,500 Ω, as a resistance that can be commercially made.

b. At the zone temperature of 27°C, the circuit above provides an emf (again from **Equation (3.24)**):

$$emf_{\text{comp}} = \frac{10}{97{,}500 + 100(1 + 0.00385 \times 27)} 100(1 + 0.00385 \times 27) - \frac{10}{97{,}500 + 100} 100$$

$$= 1.063857\ \text{mV}.$$

The term $\alpha_{\text{ba}}T$ (see **Equation (3.19)**) is

$$\alpha_{\text{ba}}T = 39.4 \times 10^{-6} \times 27 = 1.0638 \times 10^{-3}\ \text{V}.$$

The emf of the circuit including the compensation is therefore (from **Equation (3.20)**)

$$emf = 39.4 \times 10^{-6} \times 45 - 1.0638 \times 10^{-3} + 1.063857 \times 10^{-3} = 1.773057\ \text{mV}.$$

This value corresponds to a temperature T_2:

$$T_2 = \frac{1.773057 \times 10^{-3}}{39.4 \times 10^{-6}} = 45.00145°\text{C}.$$

The error is minute—only 0.003%.

The main source of error is the selection of the resistor R_2, since we have chosen a resistance that can be made commercially. A more exact value would reduce the error (the resistor may be replaced with an adjustable resistor or potentiometer). There are other sources of error, the most important of which is the nonlinear transfer function of the thermocouple (we shall discuss this next). In practice we should also expect the resistors themselves to have some tolerance as well as some temperature dependence, adding to errors. Overall, however, this method is very accurate and commonly used in thermocouple sensing.

3.3.1 Practical Considerations

Some of the properties of thermocouples have been discussed above. The choice of materials used to make the junctions is an important consideration that affects the output emf, temperature range, and resistance of the thermocouple. To aid in the selection of thermocouples and thermocouple materials, the thermocouple reference tables have been established and are supplied by standards organizations. There are three basic tables available. The first is called the thermoelectric series table, shown in **Table 3.4** for selected materials. Each material in this table is thermoelectrically negative with respect to all materials above it and positive with respect to all materials below it. This also indicates that the farther from each other a pair is, the larger the emf output that will be produced.

TABLE 3.4 ■ The thermoelectric series: selected elements and alloys at selected temperatures

100°C	500°C	900°C
Antimony	Chromel	Chromel
Chromel	Copper	Silver
Iron	Silver	Gold
Nichrome	Gold	Iron
Copper	Iron	90% platinum, 10% rhodium
Silver	90% platinum, 10% rhodium	Platinum
90% platinum, 10% rhodium	Platinum	Cobalt
Platinum	Cobalt	Alumel
Cobalt	Alumel	Nickel
Alumel	Nickel	Constantan
Nickel	Constantan	
Constantan		

TABLE 3.5 ■ Seebeck coefficients with respect to platinum 67

	Thermoelement type—Seebeck coefficient [μV/°C]					
Temperature [°C]	JP	JN	TP	TN, EN	KP, EP	KN
0	**17.9**	**32.5**	5.9	32.9	25.8	13.6
100	17.2	37.2	9.4	37.4	30.1	11.2
200	14.6	40.9	11.9	41.3	32.8	7.2
300	11.7	43.7	14.3	43.8	34.1	7.3
400	9.7	45.4	16.3	45.5	34.5	7.7
500	9.6	46.4		46.6	34.3	8.3
600	11.7	46.8		46.9	33.7	8.8
700	15.4	46.9		46.8	33.0	8.8
800				46.3	32.2	8.8
900				45.3	31.4	8.5
1000				44.2	30.8	8.2

The second standard table lists the Seebeck coefficients of various materials with reference to platinum 67 and of various common thermocouple types, as shown in **Tables 3.5** and **3.6**. In these tables the first material in each type (E, J, K, R, S, and T) is positive and the second is negative. In **Table 3.5** the Seebeck emf is given for the base elements of thermocouples with respect to platinum 67. For example, J-type thermocouples use iron and constantan. Thus column JP lists the Seebeck emf for iron with respect to platinum 67, whereas JN lists the emfs for constantan with respect to platinum 67. Adding the two together gives the corresponding value for the J-type thermocouple in **Table 3.6**. Thus, for example, taking the JP and JN values at 0°C in **Table 3.5** and adding them, $17.9 + 32.5 = 50.4$ μV/°C, gives the entry in the J column at 0°C in **Table 3.6**. Note also that these tables list limits on the high- or low-temperature use of elements and thermocouples and that the Seebeck coefficients vary with temperature. This means that the output of thermocouples cannot be linear, as we shall see shortly.

The third table, called the thermoelectric reference table, gives the thermoelectric emf produced by the thermocouple (in effect, this is the transfer function) for each type

TABLE 3.6 ■ Seebeck coefficients for various types of thermocouples

	Thermocouple type—Seebeck coefficient [μV/°C]					
Temperature [°C]	**E**	**J**	**K**	**R**	**S**	**T**
−200	25.1	21.9	15.3			15.7
−100	45.2	41.1	30.5			28.4
0	58.7	**50.4**	39.4	5.3	5.4	38.7
100	67.5	54.3	41.4	7.5	7.3	46.8
200	74.0	55.5	40.0	8.8	8.5	53.1
300	77.9	55.4	41.4	9.7	9.1	58.1
400	80.0	55.1	42.2	10.4	9.6	61.8
500	80.9	56.0	42.6	10.9	9.9	
600	80.7	58.5	42.5	11.3	10.2	
700	79.8	62.2	41.9	11.8	10.5	
800	78.4		41.0	12.3	10.9	
900	76.7		40.0	12.8	11.2	
1000	74.9		38.9	13.2	11.5	

of thermocouple as an *n*th-order polynomial in a range of temperatures. In fact, the tables give the coefficients of the polynomials. The standard tables provide the emf with a reference junction at 0°C. These tables ensure accurate representation of the thermocouple's output and can be used by the controller to accurately represent the temperature sensed by the thermocouple. There are in fact two tables. One provides the thermocouple output (with reference to zero temperature), whereas the second provides the temperature corresponding to an output emf. As an example of how these tables represent the transfer function consider **Table 3.7**, which shows the table entry

TABLE 3.7 ■ Standard thermoelectric reference table (transfer function) for type E thermocouples (chromel-constantan) with reference junction at 0°C.

Polynomial: $emf = \sum_{i=0}^{n} c_i T^i$ [μV]

Temperature range [°C]	−270°C to 0°C	0°C to 1000°C
C_0	0.0	0.0
C_1	$5.8665508708 \times 10^{1}$	$5.8665508708 \times 10^{1}$
C_2	$4.5410977124 \times 10^{-2}$	$4.5032275582 \times 10^{-2}$
C_3	$-7.7998048686 \times 10^{-4}$	$2.8908407212 \times 10^{-5}$
C_4	$-2.5800160843 \times 10^{-5}$	$-3.3056896652 \times 10^{-7}$
C_5	$-5.9452583057 \times 10^{-7}$	$6.5024403270 \times 10^{-10}$
C_6	$-9.3214058667 \times 10^{-9}$	$-1.9197495504 \times 10^{-13}$
C_7	$-1.0287605534 \times 10^{-10}$	$-1.2536600497 \times 10^{-15}$
C_8	$-8.0370123621 \times 10^{-13}$	$2.1489217569 \times 10^{-18}$
C_9	$-4.3979497391 \times 10^{-15}$	$-1.4388041782 \times 10^{-21}$
C_{10}	$-1.6414776355 \times 10^{-17}$	$3.5960899481 \times 10^{-25}$
C_{11}	$-3.9673619516 \times 10^{-20}$	
C_{12}	$-5.5827328721 \times 10^{-22}$	
C_{13}	$-3.4657842013 \times 10^{-26}$	

In explicit polynomial form we have:

In the range $-270°C$ to $0°C$,

$$
\begin{aligned}
emf = {} & 5.8665508708 \times 10^{1} T^{1} + 4.5410977124 \times 10^{-2} T^{2} - 7.7998048686 \times 10^{-4} T^{3} \\
& - 2.5800160843 \times 10^{-5} T^{4} - 5.9452583057 \times 10^{-7} T^{5} - 9.3214058667 \times 10^{-9} T^{6} \\
& - 1.0287605534 \times 10^{-10} T^{7} - 8.0370123621 \times 10^{-13} T^{8} - 4.3979497391 \times 10^{-15} T^{9} \\
& - 1.6414776355 \times 10^{-17} T^{10} - 3.9673619516 \times 10^{-20} T^{11} \\
& - 5.5827328721 \times 10^{-22} T^{12} - 3.4657842013 \times 10^{-26} T^{13} \quad [\mu V].
\end{aligned}
$$

In the range $0°C$ to $1000°C$,

$$
\begin{aligned}
emf = {} & 5.8665508708 \times 10^{1} T^{1} + 4.5032275582 \times 10^{-2} T^{2} + 2.8908407212 \times 10^{-5} T^{3} \\
& - 3.3056896652 \times 10^{-7} T^{4} + 6.5024403270 \times 10^{-10} T^{5} \\
& - 1.9197495504 \times 10^{-13} T^{6} - 1.2536600497 \times 10^{-15} T^{7} \\
& + 2.1489217569 \times 10^{-18} T^{8} - 1.4388041782 \times 10^{-21} T^{9} \\
& + 3.5960899481 \times 10^{-25} T^{10} \quad [\mu V].
\end{aligned}
$$

(coefficients of the polynomial) for type E thermocouples, and **Table 3.8**, which shows the coefficients of the inverse polynomial, that is, the coefficients of the polynomial that provides the temperature given the emf. **Table 3.8** also shows the accuracy expected in the various temperature ranges. Note that the temperature is given in degrees Celsius and emf is given in microvolts.

The polynomials are considered to be exact. Truncation of the polynomials should be avoided since any truncation may cause large errors.

The thermoelectric reference tables for the most common types of thermocouples are given in **Appendix B**.

TABLE 3.8 ▪ Coefficients of the inverse polynomials, type E thermocouples

$$T = \sum_{i=0}^{n} c_i E^i \quad [°C]$$

Temperature range [°C]	−200°C to 0°C	0°C to 1000°C
Voltage range [μV]	$E = -8825$ to 0	$E = 0$ to 76,373
C_0	0.0	0.0
C_1	1.6977288×10^{-2}	1.7057035×10^{-2}
C_2	$-4.3514970 \times 10^{-7}$	$-2.3301759 \times 10^{-7}$
C_3	$-1.5859697 \times 10^{-10}$	$6.5435585 \times 10^{-12}$
C_4	$-9.2502871 \times 10^{-14}$	$-7.3562749 \times 10^{-17}$
C_5	$-2.6084314 \times 10^{-17}$	$-1.7896001 \times 10^{-21}$
C_6	$-4.1360199 \times 10^{-21}$	$8.4036165 \times 10^{-26}$
C_7	$-3.4034030 \times 10^{-25}$	$-1.3735879 \times 10^{-30}$
C_8	$-1.1564890 \times 10^{-21}$	$1.0629823 \times 10^{-35}$
C_9		$-3.2447087 \times 10^{-41}$
Error range	0.04°C to −0.01°C	0.02°C to −0.02°C

In explicit form we get (E denotes the thermocouple emf in μV):

In the range −270°C to 0°C,

$$T = 1.6977288 \times 10^{-2}E^1 - 4.3514970 \times 10^{-7}E^2 - 1.5859697 \times 10^{-10}E^3 - 9.2502871 \times 10^{-14}E^4 - 2.6084314 \times 10^{-17}E^5 - 4.1360199 \times 10^{-21}E^6 - 3.4034030 \times 10^{-25}E^7 - 1.1564890 \times 10^{-21}E^8 \quad [°C].$$

In the range 0°C to 1000°C,

$$T = 1.7057035 \times 10^{-2}E^1 - 2.3301759 \times 10^{-7}E^2 + 6.5435585 \times 10^{-12}E^3 - 7.3562749 \times 10^{-17}E^4 - 1.7896001 \times 10^{-21}E^5 + 8.4036165 \times 10^{-26}E^6 - 1.3735879 \times 10^{-30}E^7 + 1.0629823 \times 10^{-35}E^8 - 3.2447087 \times 10^{-41}E^9 \quad [°C].$$

EXAMPLE 3.9 Thermoelectric reference tables

A chromel-constantan thermocouple is intended for use in a steam generator, normally operating at 350°C. To measure the temperature it is suggested to use a reference junction at 100°C (boiling water) since that is easier to maintain in the steam plant. In addition, to simplify interfacing, it is suggested to use only the first three terms in the polynomial for the reference emf.

a. Calculate the thermoelectric emf produced by the thermocouple at the nominal temperature (350°C).
b. Calculate the error incurred by using only the first three terms in the polynomial.

Solution:

a. The polynomial in **Table 3.7** gives the output for the chromel-constantan thermocouple (E-type) with a 0°C reference junction. We calculate the output using the first three terms and subtract from it the emf of the reference junction, also calculated with the three-terms polynomial.

The emf of the thermocouple using three terms in the polynomial is

$$emf(0) = (5.8695857799 \times 10 \times T + 4.3110945462 \times 10^{-2} \times T^2 + 5.7220358202 \times 10^{-5} \times T^3) \times 10^3,$$

where $T = 350$°C. This gives 28.278 mV.

Using the same relation, the reference emf at 100°C is 6.3579 mV. Thus the measured emf is

$$emf = 28.278 - 6.3579 = 21.9201 \text{ mV}.$$

b. Using the full polynomial in **Table 3.8**, we get

$$emf(0) = 24.9614 \text{ mV}.$$

The reference emf (at 100°C, also calculated with the ninth-order polynomial) is

$$emf_{\mathrm{ref}} = 6.3171 \text{ mV}.$$

The emf is therefore

$$emf = 24.9614 - 6.3171 = 18.6443 \text{ mV}.$$

This produces an error of

$$error = \frac{21.9201 - 18.6443}{18.6443} \times 100 = 17.6\%.$$

Note that the error is due to the incomplete polynomial and has nothing to do with the reference temperature. Nevertheless, using a zero reference temperature would produce a larger output with a lower error.

Using the results in (a) and (b) for the 0°C reference values, the error would be

$$error = \frac{28.278 - 24.9614}{24.9614} \times 100 = 13.3\%.$$

TABLE 3.9 ■ Common thermocouple types and some of their properties

Materials	Sensitivity [μV/°C] at 25°C	Standard type designation	Recommended temperature range [°C]	Notes
Copper/constantan	40.9	T	0 to 400 (−270 to 400)	Cu/60% Cu 40% Ni
Iron/constantan	51,7	J	0 to 760 (−210 to 1200)	Fe/60% Cu 40% Ni
Chromel/alumel	40.6	K	−200 to 1300 (−270 to 1372)	90% Ni 10% Cr/95% Ni 2% Al 2% Mg 1% Si
Chromel/constantan	60.9	E	−200 to 900 (−270 to 1000)	90% Ni 10% Cr/60% Cu 40% Ni
Platinum (10%)/ rhodium-platinum	6.0	S	0 to 1450 (−50 to 1760)	Pt/90% Pt 10% Rh
Platinum (13%)/ rhodium-platinum	6.0	R	0 to 1600 (−50 to 1760)	Pt/87% Pt 13% Rh
Silver/palladium	10		200 to 600	
Constantan/tungsten	42.1		0 to 800	
Silicon/aluminum	446		−40 to 150	
Carbon/silicon carbide	170		0 to 2000	
Platinum (30%)/ rhodium-platinum	6.0	B	0 to 1820	Pt/70% Pt 30% Rh
Nickel/chromium-silicon alloy		N	(−270 to 1260)	
Tungsten 5%-rhenium/ tungsten 26%-rhenium		C	0 to 2320	
Nickel-18% molybdenum/ nickel-0.8% cobalt		M	−270 to 1000	
Chromel-gold/iron	15		1.2 to 300	

The common thermocouple types are shown in **Table 3.9**, which lists their basic range and transfer functions together with some additional properties. There are many other thermocouples available commercially and still more that can be made. Two chromel/ alumel thermocouples (K-type) with exposed junctions are shown in **Figure 3.12**.

FIGURE 3.12 ■ Chromel-alumel (K-type) thermocouples showing the junction.

Note: The temperature ranges shown are recommended. Nominal ranges are shown in parentheses and are higher than the recommended ranges. The sensitivity of a thermocouple is the relative Seebeck coefficient of the combination of two materials used for the thermocouple (see **Table 3.3**).

EXAMPLE 3.10 Errors in the use of thermocouples

Thermocouples must be handled carefully and connections must be properly done or significant errors will occur in the measured output. To understand this, consider a type K thermocouple (chromel-alumel) used to measure the temperature of glass in a furnace in a glass-blowing studio. The temperature needed for proper blowing is 900°C. The thermoelectric voltage is measured using the configuration in **Figure 3.13a** with a chromel-alumel reference junction at 0°C. In the connection process, the wires for the reference junction have been inadvertently inverted and now the configuration is as in **Figure 3.13b**. The junction box is at the ambient temperature of 30°C.

a. Calculate the error in the measured voltage due to the error in connection.

b. What is the temperature that the measuring instrument will show?

Solution:

a. The two connections in the temperature zone in **Figure 3.13b** are in fact two K-type thermocouples with polarity opposing the polarity of the measuring thermocouple as indicated in the figure. This has the net effect of reducing the output emf and hence showing a lower temperature.

To calculate the emf, we use the polynomial for the K-type thermocouple in **Appendix B, Section B.2**. The emf of the sensing junction is

$$\begin{aligned} E = &-1.7600413686 \times 10^{1} + 3.8921204975 \times 10^{1} \times 900 + 1.8558770032 \\ &\times 10^{-2} \times 900^{2} - 9.9457592874 \times 10^{-5} \times 900^{3} + 3.1840945719 \times 10^{-7} \\ &\times 900^{4} - 5.6072844889 \times 10^{-10} \times 900^{5} + 5.6075059059 \times 10^{-13} \\ &\times 900^{6} - 3.2020720003 \times 10^{-16} \times 900^{7} + 9.7151147152 \times 10^{-20} \times 900^{8} \\ &- 1.2104721275 \times 10^{-23} \times 900^{9} + 1.185976 \times 10^{2} e^{-1.183432\times10^{-4}(900-126.9686)^{2}} \\ = &\ 37{,}325.915\ \mu\text{V}. \end{aligned}$$

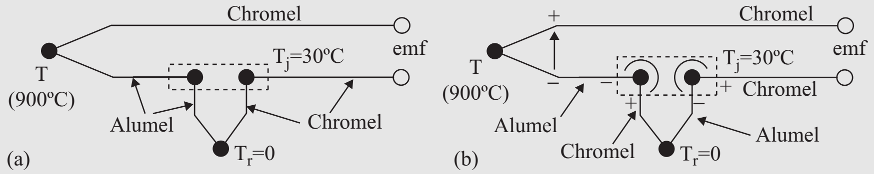


FIGURE 3.13 ■ (a) A properly connected reference junction in a chromel-alumel thermocouple. (b) The reference junction with inverted connections.

The emf of each of the two connections in the temperature zone is

$$\begin{aligned}E = &-1.7600413686 \times 10^1 + 3.8921204975 \times 10^1 \times 30 + 1.8558770032 \\ &\times 10^{-2} \times 30^2 - 9.9457592874 \times 10^{-5} \times 30^3 + 3.1840945719 \times 10^{-7} \\ &\times 30^4 - 5.6072844889 \times 10^{-10} \times 30^5 + 5.6075059059 \times 10^{-13} \times 30^6 \\ &- 3.2020720003 \times 10^{-16} \times 30^7 + 9.7151147152 \times 10^{-20} \times 30^8 \\ &- 1.2104721275 \times 10^{-23} \times 30^9 + 1.185976 \times 10^2 e^{-1.183432\times 10^{-4}(30-126.9686)^2} \\ = &\ 1203.275\ \mu\text{V}.\end{aligned}$$

The net emf at the instrument is

$$emf = emf(900) - 2 \times emf(30) = 37.3259 - 2 \times 1.2033 = 34.9193 \text{ mV}.$$

The error is the difference between the correct value and the actual reading, or

$$error = 37.3259 - 34.9193 = 2.4066 \text{ mV}.$$

This error is due to the two reversed connections, each contributing 1.2033 mV.

b. To find the temperature that corresponds to this reading we use the inverse polynomial and substitute $E = 34{,}919.3\ \mu\text{V}$:

$$\begin{aligned}T = &-1.318058 \times 10^2 + 4.830222 \times 10^{-2} \times 34919.3 - 1.646031 \times 10^{-6} \\ &\times 34919.3^2 + 5.464731 \times 10^{-11} \times 34919.3^3 - 9.650715 \times 10^{-16} \times 34919.3^4 \\ &+ 8.802193 \times 10^{-21} \times 34919.3^5 - 3.110810 \times 10^{-26} \times 34919.3^6 \\ = &\ 839.97^\circ\text{C}.\end{aligned}$$

This represents an error of 6.67% in the temperature reading.

3.3.2 Semiconductor Thermocouples

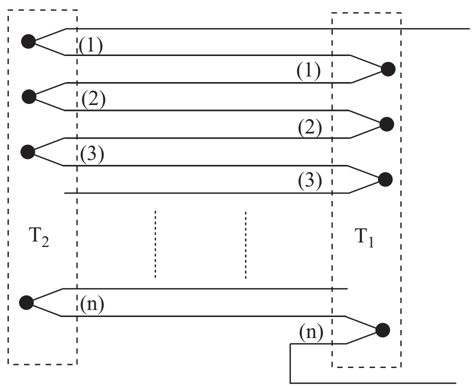
As can be seen from **Table 3.2**, semiconductors such as *p* and *n* silicon (*p*-doped and *n*-doped silicon) have absolute Seebeck coefficients that are orders of magnitude higher than those of conductors. The advantage of using semiconductors lies primarily in the large emf that develops at the junction of an *n* or *p* semiconductor and a metal (typically aluminum) or at the junction between an *n* and a *p* material. In addition, these junctions can be produced by standard semiconductor fabrication techniques, adding to their widespread use in integrated electronics. They suffer from one major shortcoming—the range of temperatures at which they are useful is limited. Silicon in general cannot operate below −55°C or above about 150°C. However, there are semiconductors such as bismuth telluride (Bi_2Te_3) that extend the range to about 225°C, and newer materials can reach about 800°C. Most semiconductor thermocouples are used either in thermopiles for sensing or in thermopiles designed for cooling and heating (Peltier cells). The latter are viewed here as actuators since their main purpose is generation of power or cooling/heating. Peltier cells can be used as sensors since their output is directly proportional to the temperature gradient across the cells. As far as their properties and use are concerned, they are similar to other semiconductor thermocouples.

3.3.3 Thermopiles and Thermoelectric Generators

A thermopile is an arrangement of a number of thermocouples so that their emfs are connected in series. The purpose of this arrangement is to provide much higher outputs than are possible with single junctions. An arrangement of this type is shown in **Figure 3.14**. Note that whereas the electrical outputs are in series, the thermal inputs are in parallel (all cold junctions are at one temperature and all hot junctions are at a second temperature). If the output of a single junction pair is emf_1, and there are n pairs in the thermopile, the output of the thermopile is $n \times emf_1$. The use of thermopiles dates back to the end of the last century and they are commonly used today for a variety of applications. In particular, semiconductor thermocouples are easily produced and integrated with electronics to form the basis for advanced integrated sensors. We will discuss these again in the following chapter where thermopiles are used in infrared sensors. Metal-based thermopiles are used both as sensors and as electricity generators. A very common thermopile sensor used in gas furnaces to detect the pilot flame has a nominal output of 750 mV (0.75 V) and uses a few dozen thermocouples to operate at temperatures up to about 800°C (see **Example 3.11**). Other thermopile assemblies are used in gas-fired generators for the purpose of generating electricity for small, remote installations.

Semiconductor thermopiles made of crystalline semiconductor material such as bismuth telluride (Bi_2Te_3) are being used in Peltier cells for cooling and heating in dual-purpose refrigerators/heaters, primarily for outdoor use and transportation of medical materials. These can also be used as sensors and can have output voltages of a few volts. In this type of semiconductor thermopile, the base semiconductor is doped to make a junction between an n- and p-type material and processed to yield oriented polycrystalline semiconductors with anisotropic thermoelectric properties. Because the junctions are small, hundreds of pairs may be built into a single device to produce outputs on the order of 20 V or more.

FIGURE 3.14 ■ Principle of a thermopile.



(a) (b)

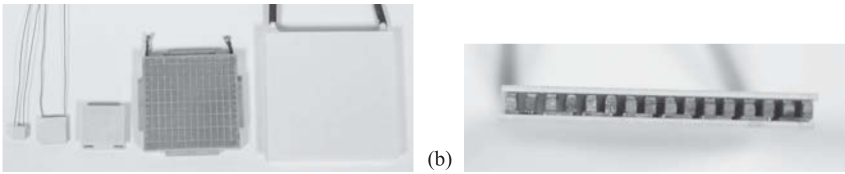


FIGURE 3.15 ■ (a) Various Peltier cells. (b) Detail of the construction of a Peltier cell.

Figure 3.15a shows a number of thermoelectric devices (Peltier cells) designed primarily for cooling electronics components such as computer processors. **Figure 3.15b** shows the internal construction where the cold junctions are connected thermally to one ceramic plate and the hot junctions are connected to the opposite plate. The junctions are connected in series in rows. The number of junction pairs can be quite high—typically 31, 63, 127, 255, etc. (the odd number allows space for connection of the lead wires in a matrix of junctions which is usually $n \times n$). A typical cooling cell operating at 12 V (or generating 12 V nominal) will contain 127 junctions. Reversing the current in a cooling device will produce heating.

EXAMPLE 3.11 Thermoelectric furnace pilot sensor

A thermopile is needed to sense the presence of a pilot light in a gas furnace to ensure that the gas valve is not opened in the absence of the pilot light. The thermopile needs to provide thermoelectric voltage (emf) of 750 mV for a temperature difference of 650°C. The cold junction is supplied by the body of the furnace, which is at 30°C.

a. What are the options for thermocouples one can use for this purpose? Select an appropriate thermocouple.
b. For the selection in (a), how many thermocouples are needed?
c. Can one use a Peltier cell for this purpose?

Solution:

a. Many of the thermocouple types can be used with the exception of type T, semiconductor thermocouples, and some like the constantan-tungsten thermocouples. A K-type (chromel-alumel), a J-type (iron-constantan), or an E-type (chromel-constantan) should work well. We will select the E-type thermocouple to build the thermopile because it produces higher emfs and thus fewer thermocouples will be needed.

b. The emf of an individual thermocouple is calculated using the coefficients in **Table 3.7** with a reference temperature of 30°C as follows:

At 650°C,

$$\begin{aligned} emf &= 5.8665508708 \times 10^{1} \times 650 + 4.5032275582 \times 10^{-2} \times 650^{2} + 2.8908407212 \\ &\quad \times 10^{-5} \times 650^{3} - 3.3056896652 \times 10^{-7} \times 650^{4} + 6.5024403270 \times 10^{-10} \\ &\quad \times 650^{5} - 1.9197495504 \times 10^{-13} \times 650^{6} - 1.2536600497 \times 10^{-15} \times 650^{7} \\ &\quad + 2.1489217569 \times 10^{-18} \times 650^{8} - 1.4388041782 \times 10^{-21} \times 650^{9} \\ &\quad + 3.5960899481 \times 10^{-25} \times 650^{10} \\ &= 49{,}225.67\ \mu\text{V}. \end{aligned}$$

At 30°C,

$$\begin{aligned} emf &= 5.8665508708 \times 10^{1} \times 30 + 4.5032275582 \times 10^{-2} \times 30^{2} + 2.8908407212 \\ &\quad \times 10^{-5} \times 30^{3} - 3.3056896652 \times 10^{-7} \times 30^{4} + 6.5024403270 \times 10^{-10} \\ &\quad \times 30^{5} - 1.9197495504 \times 10^{-13} \times 30^{6} - 1.2536600497 \times 10^{-15} \times 30^{7} \\ &\quad + 2.1489217569 \times 10^{-18} \times 30^{8} - 1.4388041782 \times 10^{-21} \times 30^{9} \\ &\quad + 3.5960899481 \times 10^{-25} \times 30^{10} \\ &= 1801.022\ \mu\text{V}. \end{aligned}$$

The emf at 650°C with respect to zero is 49.225 mV. The emf at 30°C with respect to zero is 1.801 mV. Therefore the emf at 650°C with reference to 30°C is 49.225−1.801 = 47.424 mV.

For a 750 mV output one needs

$$n = \frac{750}{47.424} = 15.8 \rightarrow n = 16.$$

c. No and perhaps yes. Most Peltier cells are based on low-temperature semiconductors and thus cannot be used directly. However, there are high-temperature Peltier cells that may be used, but even in the absence of these, one can place the Peltier cell with the cold junction on the furnace body and provide a metal structure to conduct heat from the pilot light to the hot surface of the Peltier cell while ensuring that the temperature on the hot surface does not exceed about 80°C (most Peltier cells operate at a temperature difference below 50°C between the hot and cold surfaces). The advantage of the Peltier cell is its physical size for a given thermoelectric voltage and, of course, the fact that it can generate higher emfs than metal thermopiles.

3.4 | *p-n* JUNCTION TEMPERATURE SENSORS

Returning now to semiconductors, suppose that an intrinsic semiconductor is doped so that part of it is of *p*-type while the other is of *n*-type, as shown in **Figure 3.16a**. By so doing a *p-n* junction is created. This is usually indicated as shown in **Figure 3.16b** and is known as a diode. The direction of the arrow shows the direction of flow of current (holes). Electrons flow in the opposite direction. The diode is said to conduct when forward biased, as shown in **Figure 3.16c**. When reverse biased (**Figure 3.16d**), the diode does not conduct. The current–voltage (I-V) characteristics of a *p-n* junction are shown in **Figure 3.17**.

When a *p-n* junction is forward biased, the current through the diode is temperature dependent. This current can be measured and used to indicate temperature. Alternatively, the voltage across the diode can be measured (almost always a preferable approach) and its dependence on temperature used as the sensor's output. This type of sensor is called a *p-n* junction temperature sensor or bandgap temperature sensor. It is particularly useful because it can be easily integrated in microcircuits and is rather linear in output. Needless to say, any diode or a junction in a transistor can be used for this purpose.

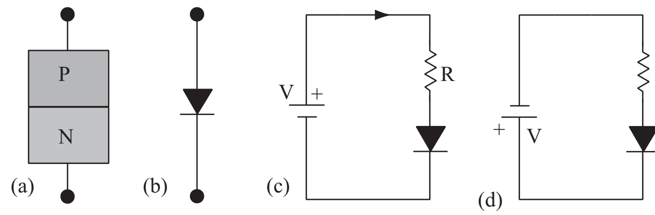


FIGURE 3.16 ■ (a) Schematic of a *p-n* junction. (b) The symbol of the junction as a diode. (c) Forward biasing of the diode. (d) Reverse biasing of the diode.

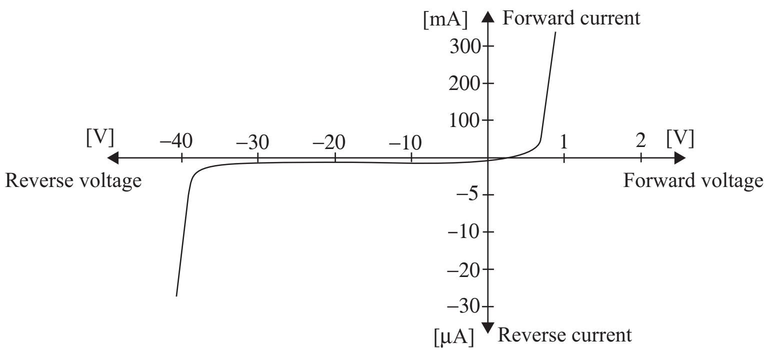


FIGURE 3.17 ■ The *I-V* characteristics of a silicon diode.

Assuming that the junction is forward biased, the *I-V* characteristic is described by the following relation:

$$I = I_0\left(e^{qV/nkT} - 1\right) \quad [\mathrm{A}], \tag{3.25}$$

where I_0 is the saturation current (a small current, on the order of a few nanoamps, dominated by temperature effects), q is the charge of the electron, k is Boltzmann's constant, and T is the absolute temperature (K). n is a constant between 1 and 2 depending on a number of properties including the materials involved and may be viewed as a property of the device. In junction temperature sensors, the current I is reasonably large compared with I_0, so that the term -1 can be neglected. Also, $n = 2$ for this type of sensor, so a good approximation to the forward current in a diode is

$$I \approx I_0 e^{qV/2kT} \quad [\mathrm{A}]. \tag{3.26}$$

Even if the terms affecting the *p-n* junction characteristics are not known, a junction can always be calibrated to measure temperature in a given range. The relation between current and temperature is nonlinear, as can be seen in **Equation (3.26)**. Usually the voltage across the diode is both easier to sense and more linear. The latter is given as

$$V_f = \frac{E_g}{q} - \frac{2kT}{q}\ln\left(\frac{\mathrm{C}}{I}\right) \quad [\mathrm{V}], \tag{3.27}$$

where E_g is the bandgap energy (in joules) of the material (to be discussed more fully in **Chapter 4**; see **Table 4.3** for specific values), C is a temperature-independent constant for the diode, and I is the current through the junction. If the current is constant, then the voltage is a linear function of temperature with negative slope. The slope (dV/dT) is clearly current dependent and varies with the semiconductor material. For silicon it is between 1.0 and 2.5 mV/°C depending on the current. This is shown in **Figure 3.18** for silicon diodes in the range −50°C to +150°C. The voltage across the diode at room temperature is approximately 0.7 V for silicon diodes (the larger the current through the diode, the higher the forward voltage drop across the diode at any given temperature). **Equation (3.27)** may be used to design a sensor based on almost any available diode or transistor. In general the published values for the bandgap energy for silicon can be used and the constant C determined by measuring the forward voltage drop at a given temperature and current through the diode.

In using a diode as a temperature sensor, a stable current source is needed. In most practical applications a voltage source with a relatively large resistor may be used, as in **Figure 3.19**, to bias the junction and establish a small current on the order of 100–200 μA. Because V_f is not constant with temperature, this method of biasing is only sufficient for general purpose applications or when the sensing span is small. Sensitivities are between 1 and 10 mV/°C with a thermal time response of less than 1 second. Like any thermal sensor, self-heating must be taken into account when connecting power to junction sensors. The self-heating effects are similar to those of thermistors and are on the order of 0.1–1 mW/°C. More sophisticated methods of biasing will be discussed in **Chapter 11** when we take up the issue of current sources. Junction sensors can be fabricated on a silicon chip together with all accompanying components, including current regulation, and may be quite complicated. The sensitivity of these junction sensors is usually improved to about 10 mV/°C and may be calibrated to produce output proportional to temperature on the Celsius, Fahrenheit, or Kelvin scales. The device is usually connected to a constant voltage source (say 5 V) and produces an output voltage directly proportional to the selected scale, with excellent linearity and

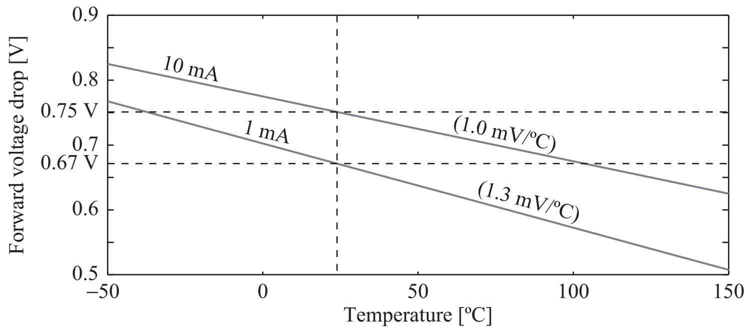


FIGURE 3.18 ■ Potential drop on a forward-biased *p-n* junction versus temperature (1N4148 silicon switching diode, evaluated experimentally).

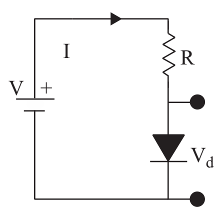


FIGURE 3.19 ■ Forward-biased *p-n* junction with a rudimentary current *source*. *R* must be large to produce a low current and small variations with variations in *V*.

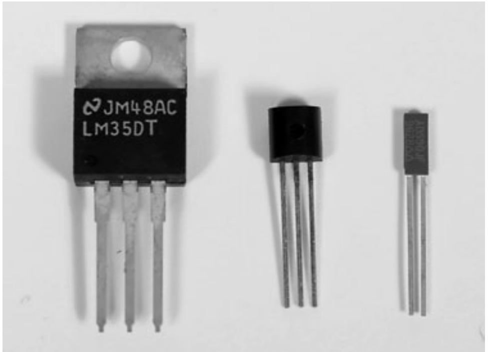


FIGURE 3.20 ■ Junction temperature sensors.

accuracy typically around ±0.1°C. The range of temperatures that can be measured with these sensors is rather small and cannot exceed the operating range of the base material. Typically a silicon sensor can operate between −55°C and about 150°C, although they can be designed for somewhat wider ranges, whereas some devices are rated for narrower ranges (usually at lower cost). **Figure 3.20** shows three junction temperature sensors in various packages.

EXAMPLE 3.12 Silicon diode as a temperature sensor

A silicon diode is proposed for use as a temperature sensor in a vehicle to sense ambient temperature between −45°C and +45°C. To determine its response, the diode is forward biased with a 1 mA current and its forward voltage drop is measured at 0°C as 0.712 V. The bandgap energy of silicon is 1.11 eV. Calculate

a. The output expected for the span needed.

b. The sensitivity of the sensor.

c. The error in measuring temperature if the self-heating specification for the diode in still air is 220°C/W.

Solution: We first calculate the constant C in **Equation (3.27)** from the known voltage drop at 0°C followed by the forward voltage drop at the range points of the span.

a. At 0°C we have

$$V_f = \frac{E_g}{q} - \frac{2kT}{q}\ln\left(\frac{C}{I}\right) = 0.712$$

$$= \frac{1.11 \times 1.602 \times 10^{-19}}{1.602 \times 10^{-19}} - \frac{2 \times 1.38 \times 10^{-23} \times 273.15}{1.602 \times 10^{-19}}\ln\left(\frac{C}{10^{-3}}\right).$$

Note that 1 eV $= 1.602 \times 10^{-19}$ J and that the temperature T is in degrees kelvin. Thus

$$0.712 = 1.11 - 0.04706\ln\left(\frac{C}{10^{-3}}\right) \rightarrow \ln\left(10^3 C\right) = \frac{0.712 - 1.11}{-0.04706} = 8.457.$$

This gives

$$e^{8.457} = 10^3\mathrm{C} \rightarrow \mathrm{C} = \frac{e^{8.457}}{10^3} = 4.7.$$

Now the forward voltage at $-45°\mathrm{C}$ and $+45°\mathrm{C}$ are

$$\begin{aligned} V_f(-45°\mathrm{C}) &= \frac{E_g}{q} - \frac{2kT}{q}\ln\left(\frac{\mathrm{C}}{I}\right) \\ &= \frac{1.11 \times 1.602 \times 10^{-19}}{1.602 \times 10^{-19}} - \frac{2 \times 1.38 \times 10^{-23} \times (273.15 - 45)}{1.602 \times 10^{-19}}\ln\left(4.7 \times 10^3\right) \\ &= 0.77765\ \mathrm{V} \end{aligned}$$

and

$$\begin{aligned} V_f(+45°\mathrm{C}) &= \frac{E_g}{q} - \frac{2kT}{q}\ln\left(\frac{\mathrm{C}}{I}\right) = 1.11 - \frac{2 \times 1.38 \times 10^{-23} \times (273.15 + 45)}{1.602 \times 10^{-19}}\ln\left(4.7 \times 10^3\right) \\ &= 0.64654\ \mathrm{V}. \end{aligned}$$

The forward voltage drop varies between 0.77765 V at $-45°\mathrm{C}$ and 0.64654 V at $+45°\mathrm{C}$.

b. Since **Equation (3.27)** is linear with temperature, the sensitivity of the device is the difference between the two range points divided by the difference in temperature:

$$s = \frac{0.64654 - 0.77765}{90} = 1.457\ \mathrm{mV}/°\mathrm{C}.$$

Comparing this to **Figure 3.18**, it is clear that the diode selected here is somewhat less sensitive than the one described in **Figure 3.18**.

c. The self-heating effect causes an increase in temperature of 220°C/W or 0.22°C/mW. Since the current through the diode is 1 mA, the power dissipated is,

at $-45°\mathrm{C}$,

$$P(-45°\mathrm{C}) = 0.77765 \times 10^{-3} = 0.778\ \mathrm{mW}.$$

The increase in temperature is $0.778 \times 0.22 = 0.171°\mathrm{C}$ and the forward voltage decreases by $0.171 \times 1.457 = 0.249$ mV. This represents an error of 0.38% in the temperature reading.

At 45°C,

$$P(+45°\mathrm{C}) = 0.64654 \times 10^{-3} = 0.647\ \mathrm{mW}.$$

The increase in temperature is $0.647 \times 0.22 = 0.142°\mathrm{C}$ and the forward voltage decreases by $0.142 \times 1.457 = 0.207$ mV. The error in temperature reading is 0.32%.

These errors are small, but not necessarily negligible.

3.5 | OTHER TEMPERATURE SENSORS

Almost all physical quantities and phenomena that can be measured are temperature dependent and therefore, in principle, a sensor can be designed around almost any of these. For example, the speed of light and/or its phase in an optical fiber, the speed of sound in air or a fluid, the frequency of vibration of a piezoelectric membrane, the length of a piece of metal, the volume of a gas, and so on, are temperature dependent. Rather than discuss all possible sensors of this type, we briefly discuss here a few representative sensors.

3.5.1 Optical and Acoustical Sensors

Optical temperature sensors are of two basic types. One type is the noncontact sensors that measure the infrared radiation of a source. With proper calibration the temperature of the source can be sensed and accurately measured. We shall discuss infrared radiation sensors in the **Chapter 4.** There are, however, many other temperature sensors based on the optical properties of materials. For example, the index of refraction of silicon is temperature dependent and the speed of light through a medium is inversely proportional to the index of refraction. By comparing the phase of a beam propagating through a silicon fiber that is exposed to heat with the phase of a beam in a reference fiber, the phase can be used as a measure of temperature. This type of sensor is an interferometric sensor and can be extremely sensitive, especially if the fiber is long.

Acoustic sensors can act in a similar way, but because the speed of sound is low, it can be measured directly from the time of flight of an acoustic signal through a known distance. Typically a sensor of this type includes a source that generates an acoustic signal, such as a loudspeaker or ultrasonic transmitter (a device very similar to a loudspeaker, but much smaller and operating at higher frequencies). A tube filled with a gas or a fluid is exposed to the temperature to be measured and a microphone or a second ultrasonic device (a receiver) is placed at the other end. This is shown schematically in **Figure 3.21**. A signal is transmitted and the delay between the time of its transmission and the time of arrival at the receiver is calculated. The length of the tube divided by the time difference (time of flight) gives the speed of sound. This can then be calibrated to read the temperature. For example, if the tube is air filled, the relation between temperature and the speed of sound in air is

$$v_s = 331.5\sqrt{\frac{T}{273.15}} \quad [\mathrm{m/s}], \tag{3.28}$$

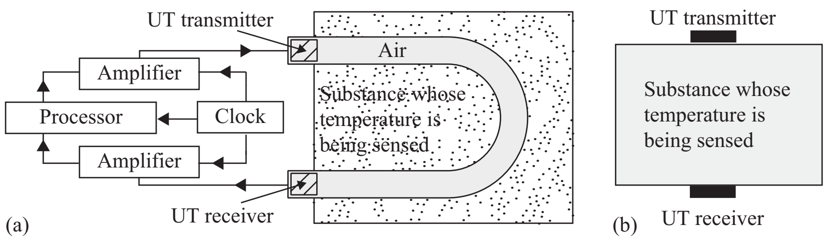


FIGURE 3.21 ■ Acoustic temperature sensing. (a) Sound travels through a fluid-filled channel. (b) Sound travels through the working fluid itself.

where T is the absolute temperature (K) and the speed at 273.15 (0°C) is 331.5 m/s. In some installations, the tube may be eliminated and the substance whose temperature is being sensed may serve instead (**Figure 3.21b**).

The speed of sound in water is also temperature dependent, and this dependency can be used to measure temperature or the temperature may be used to compensate for changes in the speed of sound due to temperature. In seawater the speed of sound depends on depth as well as salinity. Disregarding salinity and depth (i.e., regular water at the surface), a simplified relation is

$$\begin{aligned} v_s &= a + bT + cT^2 + dT^3 \quad [\text{in m/s}], a = 1449, b = 4.591, \\ c &= -5.304 \times 10^{-2}, \quad d = 2.374 \times 10^{-4}, \end{aligned} \tag{3.29}$$

where $a = 1449$ m/s is the speed of sound in water at 0°C and T is given in degrees Celsius. The other terms may be viewed as corrections to the first term. In seawater and at other depths additional correction terms are needed.

3.5.2 Thermomechanical Sensors and Actuators

An important and common class of temperature sensors and actuators is the so-called thermomechanical sensors and actuators. The general idea is that the temperature being sensed changes a physical property such as length, pressure, volume, etc. These properties then serve to measure the temperature and often perform actuation. For this reason, and since it is often difficult to distinguish between sensing and actuation, these devices will be discussed together. Common examples are the change in length of a metal or the volume of a gas. Another is the glass thermometer in which the height of a capillary column of fluid (mercury or alcohol) indicates the temperature through expansion of the fluid. Many of these types of sensors also feature a direct reading of temperature without the need for an intermediary processing stage and typically require no external power.

A simple example of a sensor based on the expansion of gas (or an expandable fluid such as alcohol) is shown in **Figure 3.22**. The volume of the gas, and therefore the position of the piston is directly proportional to the temperature being measured.

The volume expands based on the volume expansion coefficient of the medium, β. The change in volume due to the change in temperature is

$$\Delta V = \beta V \Delta T \quad [\text{m}^3], \tag{3.30}$$

where V is the volume and ΔT is the change in temperature. The coefficient β is a property of the material and is typically given in 10^{-6} per degree Celsius at a specific

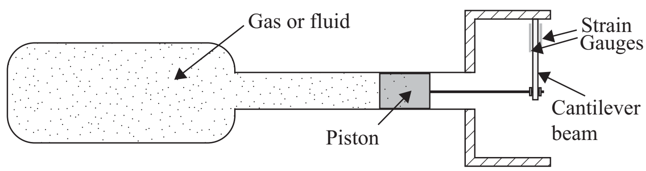


FIGURE 3.22 ■ Temperature sensing based on expansion of gases (or liquids). The piston may be replaced with a diaphragm.

temperature (usually 20°C). The volume of a material at a temperature T is then calculated as

$$V = V_0[1 + \beta(T - T_0)] \quad [\mathrm{m}^3], \tag{3.31}$$

where V_0 is the volume at the reference temperature T_0 (usually 20°C).

In many solids and fluids (isotropic materials), the relation between the volume coefficient of expansion, β, and the linear coefficient of expansion, α, is

$$\alpha = \frac{\beta}{3}. \tag{3.32}$$

If necessary, one can also define a coefficient of surface expansion as

$$\gamma = \frac{2\beta}{3}. \tag{3.33}$$

In the case of gases, the situation is somewhat different as the expansion depends on the conditions under which the gas expands and, of course, only a volume expansion is physically meaningful. For an ideal gas, under isobaric expansion (i.e., the gas pressure remains unchanged), the coefficient of volume expansion is

$$\beta = \frac{1}{T}, \tag{3.34}$$

where T is the absolute temperature.

The expansion of gasses can also be deduced from the ideal gas law, stated as follows:

$$PV = nRT, \tag{3.35}$$

where P is pressure in pascals, V is volume in cubic meters, T is temperature in degrees kelvin, n is the amount of gas in moles, and R is the gas constant, equal to 8.314 J/K/mol or 0.08206 L/atm/K/mol. This relation gives the state of the gas and may be used to calculate pressure under constant volume conditions or volume under constant pressure conditions as a function of temperature. One can then use either **Equation (3.31)** with the coefficient in **Equation (3.34)** or start with the ideal gas law in **Equation (3.35)** to calculate the volume of the gas. The choice depends on the conditions (see **Problem 3.29**).

Coefficients of linear and volume expansion are given in **Table 3.10** for a number of metals, fluids, and other materials. It should be noted that some materials have low coefficients, whereas others, particularly fluids, have large coefficients. Clearly the use of fluids like ethanol or water is favored because the resulting expansion is larger, leading to a more sensitive sensor/actuator. Of course, gases expand much more and thus a gas-filled sensor will have higher sensitivity and faster response time, but, on the other hand, the sensor will be nonlinear.

The position of the piston (**Figure 3.22**) is sensed in any of a number of ways. It may drive a potentiometer and the resistance is then a measure of temperature. Alternatively, it may have a mirror connected to it that tilts with the increase in pressure and a light beam is then deflected accordingly, or one can measure the strain in a diaphragm to get the temperature. Or it may even drive a needle directly to indicate temperature on a scale. In the configuration shown in **Figure 3.22**, the temperature is sensed using strain gauges that measure the strain in the cantilever beam (we shall discuss strain gauges in

TABLE 3.10 ■ Coefficients of linear and volume expansion for some materials given per degree Celsius at 20°C

Material	Coefficient of linear expansion (α), $\times 10^{-6}$/°C	Coefficient of volume expansion (β), $\times 10^{-6}$/°C
Aluminum	23.0	69.0
Chromium	30.0	90.0
Copper	16.6	49.8
Gold	14.2	42.6
Iron	12.0	36
Nickel	11.8	35.4
Platinum	9.0	27.0
Phosphor-bronze	9.3	27.9
Silver	19.0	57.0
Titanium	6.5	19.5
Tungsten	4.5	13.5
Zinc	35	105
Quartz	0.59	1.77
Rubber	77	231
Mercury	61	182
Water	69	207
Ethanol	250	750
Wax	16,000–66,000	50,000–200,000

Chapter 6). The piston may be replaced with an appropriate diaphragm. A very sensitive implementation of this sensor is shown in **Figure 3.23**. This sensor is called a Golay cell (sometimes also called a thermopneumatic sensor). The trapped gas (or liquid) expands the diaphragm and the position of the light beam indicates the temperature. Although gases or liquids may be used, gases have lower heat capacities (require less energy to raise their temperature) and hence have better response times.

We shall see in **Chapter 4** that this device can be used to measure infrared radiation as well. It should be obvious that the expansion and the resulting motion of the piston can also be used for actuation or, as is the case in alcohol and mercury thermometers, for direct indication of temperature (a sensor-actuator). However, most actuators based on expansion use metals and will be discussed below.

The diameter must be 0.309 mm. This is in fact a capillary tube. Because it is so thin, the alcohol is dyed (typically red or blue) and the tube is fitted with a cylindrical lens along its visible surface to facilitate reading of the temperature.

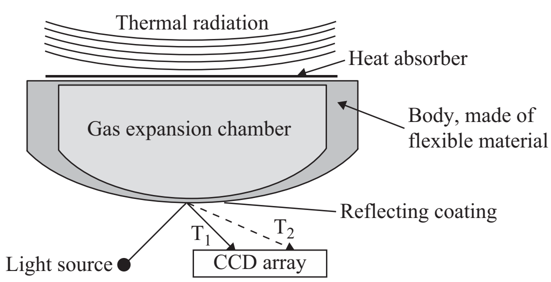


FIGURE 3.23 ■ The Golay cell is a thermopneumatic sensor based on the expansion of gasses. The charge-coupled device (CCD) array is a light sensor that will be discussed in **Chapter 4**.

EXAMPLE 3.13 The alcohol thermometer

A medical thermometer is to be manufactured using a thin glass tube, as shown in **Figure 3.24**, with a range of 34°C–43°C. The volume of the bulb (which serves as a reservoir) is 1 cm^3. To properly read the temperature, the graduations are to be 1 cm/°C (so that a change of 0.1°C raises the alcohol level by 1 mm). Assuming the glass does not expand (i.e., its coefficient of volume expansion is negligible), calculate the inner diameter of the glass tube needed to produce the thermometer.

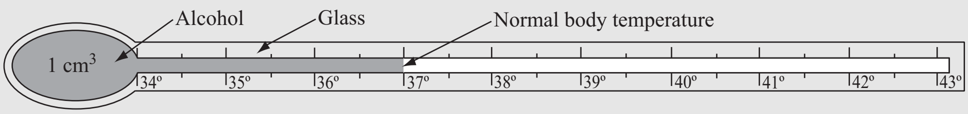


FIGURE 3.24 ■ An alcohol thermometer.

Solution: We can use **Equation (3.30)** to calculate the change in volume for a change in temperature of 9°C (43°C−34°C). This change in volume is the volume of the thin tube between the two extreme graduations (9 cm long):

$$\Delta V = \beta V \Delta T = 750 \times 10^{-6} \times 1 \times 9 = 6750 \times 10^{-6}\ \text{cm}^3,$$

where we have used the coefficient of volume expansion of ethanol (we assume, implicitly, that the coefficient is the same as at 20°C). Taking the inner diameter of the thin tube as d, we have

$$\pi \frac{d^2}{4} \times L = \Delta V \rightarrow d = \sqrt{\frac{4\Delta V}{\pi L}} \quad [\text{cm}],$$

where L is the length of the tube (9 cm in our case). Thus

$$d = \sqrt{\frac{4\Delta V}{\pi L}} = \sqrt{\frac{4 \times 6750 \times 10^{-6}}{\pi \times 9}} = 3.09 \times 10^{-2}\ \text{cm}.$$

Fluid thermometers are not as common as they used to be, but they can still be found, especially as outdoor thermometers.

Some of the oldest temperature sensors are based on the thermal expansion of metals with temperature and constitute direct conversion of temperature into displacement. These are commonly used as direct reading sensors because this mechanical expansion can be used as an actuator as well to move a dial or some other type of indicator. A conductor made in the form of a bar or wire of length l will experience an elongation with an increase in temperature. If the length is l_1 at T_1, then at $T_2 > T_1$, $l_2 > l_1$ as follows:

$$l_2 = l_1[1 + \alpha(T_2 - T_1)] \quad [\text{m}]. \tag{3.36}$$

If $T_2 < T_1$, the bar will contract. The change in length can then be measured to represent the temperature being measured. The coefficient α is called the coefficient of linear expansion of the metal (see **Table 3.10** for the coefficients for some materials). Although the coefficient of expansion is small, it is nevertheless measurable and with proper care becomes a useful method of sensing. There are two basic methods that can be

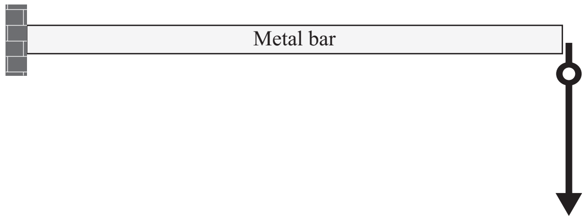


FIGURE 3.25 ■ A simple direct indication by an expanding linear bar. The dial is pushed to indicate temperature.

used. One is shown in **Figure 3.25**. It is a simple bar pushing against an indicator. An increase in temperature will move the arrow to read on a dial. An alternative way is for the bar to rotate a potentiometer or to press against a pressure gauge, in which case the electric signal may be used to indicate temperature or to connect to a processor. Although a sound principle, this is difficult to use because the expansion is small (see **Table 3.10**) and because of hysteresis and mechanical slack. However, this method is often used in microelectromechanical systems (MEMS) where expansions of a few micrometers are sufficient to affect the necessary actuation. We will discuss thermal actuation in MEMS in **Chapter 10** (but see **Example 3.14)**. One glaring exception is wax (especially paraffin), with its large coefficient. Waxes of various compositions are used in direct actuation in thermostats, especially in vehicles (see **Problems 3.32** and **3.33**).

EXAMPLE 3.14 Linear thermal microactuator

An actuator is built as a thin chromium rectangular wire as shown in **Figure 3.26a**. A current passes through the wire to heat it up. The free end serves as the actuator (i.e., it can serve to close a switch or to tilt a mirror for an optical switch; we shall encounter these applications in **Chapter 10**). If the temperature of the chromium wire can vary from 25°C to 125°C, what is the largest change in length of the actuator? The dimensions in the figure are given at 20°C.

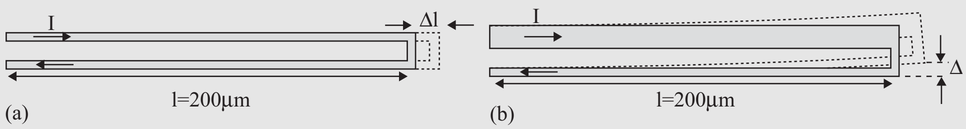


FIGURE 3.26 ■ Microthermal actuator. (a) Linear motion/actuation. (b) Angular actuation.

Solution: By direct application of **Equation (3.36)** we get

$$l(25°\mathrm{C}) = 200[1 + 30 \times 10^{-6} \times (25 - 20)] = 200.03\ \mu\mathrm{m}$$

and

$$l(125°\mathrm{C}) = 200[1 + 30 \times 10^{-6} \times (125 - 20)] = 200.63\ \mu\mathrm{m}.$$

The actuator changes in length by 0.6 μm. This may seem small, but it is both measurable and linear, and in the context of microactuators is sufficient for many applications. Thermal actuators are some of the simplest and commonly utilized microdevices. The main problem plaguing thermal actuators on the macroscopic level, that of slow response time as well as the power needed, are not relevant in microdevices. Their small size makes them sufficiently responsive and the power needed is small as well.

Note also that if one leg of the structure, say, the upper, is made thicker, it will heat to a lower temperature since it can dissipate more heat, and the whole frame will bend upwards with an increase in current. The position of the tip of the frame can then be controlled by the current in the frame (**Figure 3.26b**).

A useful implementation of the expansion of metals is the bimetal bar shown in **Figure 3.27a**. Two metals with different expansion coefficients are bonded together. Suppose that the layer on top has a higher coefficient than the layer on the bottom. When the temperature is raised, the top layer will expand more and thus the tip will move downward (**Figure 3.27b**). If the temperature is lowered, the tip will move upward. This may be used to move a dial or strain a gauge to measure the motion. It may also be used to close or open a switch. The latter is still used in direction indicators in cars, where a bimetal sensor is used to sense the current through the indicating bulb by passing the current through the bimetal material, which forms a switch as shown in **Figure 3.29a**. The bimetal bar heats up and bends downward, disconnecting the switch. The lamp goes off and the bimetal element cools down, moving upward and reconnecting the lamp. As is the case with most thermomechanical devices, the bimetal sensor is really a sensor-actuator.

The bimetal principle can also be used for direct dial indication, as shown in **Figure 3.27c**. Here a long strip of bimetal material is bent into a coil. The change in

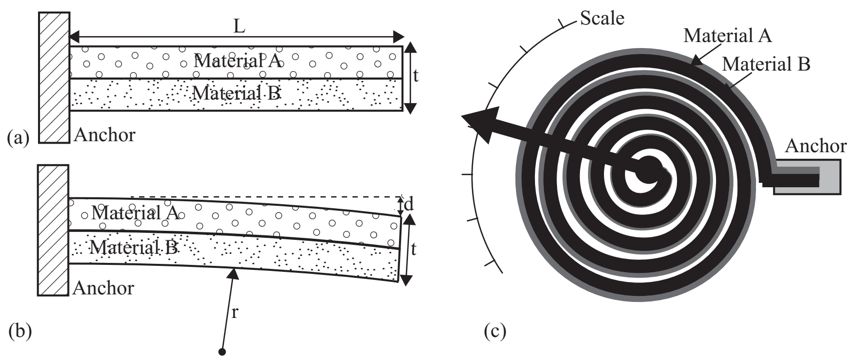
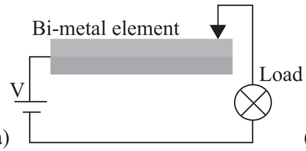
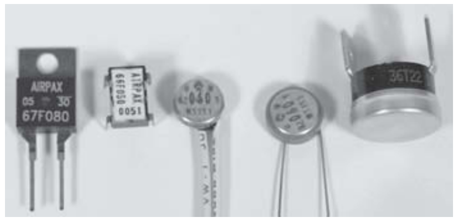


FIGURE 3.27 ■ Bimetal sensor. (a) Basic construction. (b) Displacement with temperature. (c) Coil bimetal temperature sensor.

FIGURE 3.28 ■ Bimetal thermometer and thermostat. The upper bimetal coil is the thermometer, the lower is the thermostat. The thermostat coil activates the mercury switch at the front of the picture.



(a)



(a)

FIGURE 3.29 ■ (a) Schematic of a bimetal switch as used for directional flashers in automobiles. (b) Small bimetal thermostats.

temperature causes a much larger change in the length of the strip and the motion of the strip rotates the dial in proportion to temperature.

The displacement of the free-moving end of a bimetal strip can be calculated from approximate expressions. In **Figure 3.27b**, the displacement of the free-moving tip is given by

$$d = r\left[1 - \cos\left(\frac{180L}{\pi r}\right)\right] \quad [\mathrm{m}], \tag{3.37}$$

where

$$r = \frac{2t}{3(\alpha_u - \alpha_l)(T_2 - T_1)} \quad [\mathrm{m}]. \tag{3.38}$$

T_1 is the reference temperature at which the bimetal bar is flat, T_2 is the sensed temperature, t is the thickness of the bar, L is the length of the bar, and r is the radius of curvature or warping. α_u is the coefficient of expansion of the upper conductor in the bimetal strip and α_l is the coefficient of expansion of the lower conductor.

The coil-type bimetal sensor in **Figure 3.27c** relies on the difference in expansion of the two materials to turn a dial (typically). The difference in length of the inner and outer strips causes the coil to rotate, and because the overall length is significant, the change is relatively large (see **Example 3.15**). Many simple thermometers are of this type (especially outdoor thermometers). **Figure 3.28** shows a house thermostat. The upper bimetal coil is a thermometer to indicate room temperature, whereas the lower coil is the thermostat coil. The glass bulb at the bottom is a mercury switch activated by the bimetal coil.

The type of sensors/actuators discussed here are some of the most common in consumer products because they are simple, rugged, and require no power, and at least partially because they are tried and true and there was no need to replace them. They can be found in kitchen thermometers (meat thermometers), appliances, thermostats, and outdoor thermometers.

As indicated earlier, these devices are equally useful as actuators since they convert heat into displacement. In selected cases this actuation is a natural choice, as in the example of the turn indicator switch, thermostats, and simple thermomechanical thermometers.

Whereas the bimetal principle exploits the linear expansion of metals, the volume expansion of gases and fluids can be used for sensing and for actuation, as has been exemplified by the Golay cell and the alcohol thermometer. There are also some solid materials that expand or contract significantly when a change of phase occurs. For example, water, when it freezes, expands by about 10%. One of the more interesting materials in this category is paraffin wax. When melting, its volume expands by anywhere from 5% to 20% (see **Table 3.10**) depending on the composition of the wax. More significantly from the point of view of actuation is the fact that different

EXAMPLE 3.15 Bimetal coil thermometer

An outdoor bimetal thermometer is made in the form of a coil, as shown in **Figure 3.27c**. The coil has six full turns, with the inner turn having a radius of 10 mm and the outer turn a radius of 30 mm when the thermometer is at 20°C. The bimetal strip is 0.5 mm thick and is made of chromium (outer strip) and nickel (inner strip) to resist corrosion. The thermometer is intended to operate between −45°C and +60°C. Estimate the change in angle the needle makes as the temperature changes from −45°C to +60°C.

Solution: It is not possible to calculate the change in angle exactly since this depends on additional parameters. However, we may assume that if the strip were straight, each of the conductors must expand according to **Equation (3.36)**. The difference in the expansion coefficients of the two metals forces them to coil. Therefore we will calculate the change in length of the outer strip, and that change is what causes the needle to move (the lower expansion coefficient of the inner, nickel strip is what causes curling of the strip). To be able to approximate the angle of the needle, we first calculate the length of the strip using an average radius. Then we calculate its length at −45°C and at +60°C. The difference between the two moves the strip along the inner loop.

The average radius of the coil is

$$R_{avg} = \frac{30 + 10}{2} = 20 \text{ mm}.$$

The length of the strip at the nominal temperature (20°C) is

$$L = 6\left(2\pi R_{avg}\right) = 6 \times 2\pi \times 20 \times 10^{-3} = 0.754 \text{ m}.$$

Now, at the temperature extremes we have

$$L(-45°\text{C}) = 0.754[1 + 30 \times 10^{-6} \times (-45 - 20)] = 0.75253 \text{ m}$$

and

$$L(60°\text{C}) = 0.754[1 + 30 \times 10^{-6} \times (60 - 20)] = 0.7549 \text{ m}.$$

The difference between the two is $\Delta L = 0.7549 - 0.75253 = 0.002375$ m, or 2.375 mm.

To estimate the angle the needle moves, we argue that this expansion is small and therefore the inner loop remains of the same radius. If its circumference is $2\pi r$, then $\Delta\alpha = (\Delta L/2\pi r) \times 360°$:

$$\Delta\alpha = \frac{\Delta L}{2\pi r_{inner}} \times 360 = \frac{0.002375 \times 360}{2 \times \pi \times 0.01} = 13.6°.$$

This means that the scale shown schematically in **Figure 3.27c** covers a 13.6° section.

compositions can be made to melt at specific temperatures. Furthermore, the transition is gradual, first the wax becomes soft (and expands slightly) and then melts, expanding much more. The opposite occurs when it solidifies. This introduces an inherent hysteresis in the process, which is often useful. All of these properties are exploited in thermostats for car engines. The thermostat (a car thermostat is shown in **Figure 3.30**) is essentially a cylinder and a piston, with the cylinder filled by a solid wax pellet. During operation, the piston is pushed by the melting wax to open the cooling water

FIGURE 3.30 ■ A car thermostat.

line when the engine has reached the preset temperature. The thermostat not only allows cooling, but ensures quick warming of the engine by keeping the water line closed until the proper temperature has been reached. Then it opens gradually as the temperature rises or closes gradually as the temperature drops to keep the engine coolant temperature within a narrow range of temperatures and by so doing ensures maximum engine efficiency. Thermostats come in a variety of shapes, sizes, and set temperatures. The temperature is set during production by the choice of the wax composition. Wax pellets for temperatures from about 20°C to 175°C exist for a variety of applications, not all of them automotive. See also **Problems 3.32** and **3.33**.

3.6 | PROBLEMS

Units of temperature and heat

3.1 Convert the absolute temperature (0 K) to degrees Celsius and to degrees Fahrenheit.

3.2 The calorie (cal) is a unit of energy equal to 0.239 J. How many electron volts does it represent? The electron-volt (eV) is the energy needed to move an electron (charge equal to 1.602×10^{-19} C) across a potential difference equal to 1 V.

Resistance temperature detectors

3.3 **Simple RTD.** An RTD can be built relatively easily. Consider a copper RTD made of magnet wire (copper wire insulated with a polymer). The wire is 0.1 mm thick and the nominal resistance required is 120 Ω at 20°C. Neglect the thickness of the insulating polymer.

a. How long must the wire be?

b. Assuming we wish to wind the copper wire into a single spiral winding 6 mm in diameter so that it can be enclosed in a stainless steel tube, what is the minimum length of the RTD?

c. Calculate the range of resistance of the RTD for use between −45°C and 120°C.

3.4 **Self-heat in RTDs and errors in sensing**. A platinum wire RTD enclosed in a ceramic body is designed to operate between −200°C and +600°C. Its nominal resistance at 0°C is 100 Ω and its TCR is 0.00385/°C (European curve). The sensor has a self-heat of 0.07°C/mW. The sensor is fed from a constant voltage source of 6 V through a fixed 100 Ω resistor and the voltage across the sensor is measured directly. Calculate the error in temperature sensed in the range 0°C−100°C. Plot the error as a function of temperature. What is the maximum error and at what temperature does it occur? Explain.

3.5 **Temperature sensing in a light bulb.** Incandescent lightbulbs use a tungsten wire as the light-producing filament by heating it to a temperature at which it is bright enough to produce light. The temperature of the wire can be estimated directly from the power rating and the resistance of the wire when it is cold. Given a 120 V, 100 W lightbulb with a resistance of 22 Ω at room temperature (20°C):

a. Calculate the temperature of the filament when the lightbulb is lit.

b. What are the possible sources of error in this type of indirect sensing? Explain.

3.6 **Accurate representation of resistance of RTDs.** The Callendar–Van Dusen polynomials **(Equations (3.6)** and **(3.7)**) can be used either with published data from common RTD materials or the coefficients of the polynomial can be determined from measurements. Suppose one decides to introduce a new line of RTDs made of nichrome (a nickel-chromium alloy) for the range between −200°C and +900°C. To evaluate the behavior of the new type of sensors, one must determine the constants a, b, and c in **Equations (3.6)** and **(3.7)**. There are common calibration points that guarantee exact known temperatures at which the resistance is measured. The common calibration points are the oxygen point (−182.962°C, equilibrium between liquid oxygen and its vapor), the triple point of water (0.01°C, the point of equilibrium temperature between ice, liquid water, and water vapor), the steam point (100°C, equilibrium point between water and its vapor), the zinc point (419.58°C, equilibrium point between solid and liquid zinc), the silver point (961.93°C), and the gold point (1064.43°C), as well as others. By selecting the appropriate temperature points and measuring the resistance at those points one can determine the coefficients. The resistance measurements for the new type of RTD are as follows: $R = 9$ Ω at the oxygen point, $R = 38$ Ω at 0°C, $R = 52$ Ω at the steam point, and $R = 98$ Ω at the zinc point. The TCR for the nichrome alloy used here is 0.004/°C.

a. Find the coefficients of the polynomials.

b. Find the resistance at −150°C and at 800°C. Compare the results with those obtained using **Equation (3.5)**. What are the errors?

3.7 **Effect of temperature gradient on accuracy.** Wire RTDs are often relatively long sensors and the temperature gradient on the sensor itself may be of concern in certain applications or in dynamic situations where the temperature changes quickly. To understand the effect, consider a platinum wire RTD of length 10 cm and a nominal resistance of 120 Ω at 20°C.

a. Calculate the temperature reading at 80°C if one end is at that temperature while the other is 1°C lower and the distribution within the sensor is linear.

b. Calculate the temperature reading as in (a) but for a parabolic distribution in which the temperature of the center of the sensor is at 79.25°C.

3.8 Indirect temperature sensing: pyrometric temperature sensor. A relatively old and well-established method of sensing high temperatures, especially that of molten metals and molten glass, is the use of color comparison. The premise is that if the color of the molten material and that of a control heated filament are the same, their temperatures must also be the same. With proper selection of comparison filaments the method can be very accurate, and it is an entirely noncontact method of sensing. A sensor of this type uses a lamp, heated through a variable resistor, and the voltage and current are read as shown in **Figure 3.31**.

a. Given a reading of V and I for a color match and given the resistance of the filament as R_0 at 20°C, calculate the temperature sensed. Take the temperature coefficient of resistance as α.

b. In an actual sensor, the filament is made of tungsten and has a resistance of 1.2 Ω at 20°C. In a particular application, the voltage across the lamp is measured as 4.85 V with a current of 500 mA. What is the temperature being sensed?

c. Discuss the possible errors involved in this type of measurement.

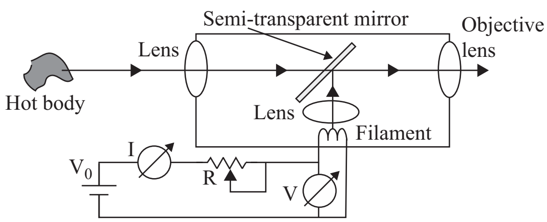


FIGURE 3.31 ■ The pyrometric temperature sensor.

3.9 TCR and its dependence on temperature. The temperature coefficient of resistance is not constant but depends on the temperature at which it is given or evaluated. Nevertheless, the formula in **Equation (3.4)** is correct at any temperature, no matter what the temperature T_0 is, as long as α is measured (or given) at T_0. Suppose α is measured at 0°C and is equal to $\alpha = 0.00385/°C$ (for platinum).

a. Calculate the coefficient α at 50°C.

b. Generalize the result in (a) as follows: given α_0 at T_0, what is α_1 at T_1?

Silicon resistive sensors

3.10 Semiconducting resistive sensor. A semiconducting resistive sensor is made as a simple rectangular bar 2 mm × 0.1 mm in cross section and 4 mm long. The intrinsic carrier concentration at 20°C is $1.5 \times 10^{10}/cm^3$ and the mobilities of electrons and holes are 1350 $cm^2/V/s$ and 450 $cm^2/V/s$, respectively. The TCR for the particular device being used here is −0.009/°C and is assumed to be unaffected by doping. Treat the sensor as a simple resistor.

a. If intrinsic material is used, calculate the resistance of the sensor at 50°C.

b. Now the material is heavily doped with an n-type dopant at a concentration of $10^{15}/m^3$. Calculate the resistance of the sensor at 50°C.

c. What is the resistance of the sensor if instead it is doped with a p-type dopant at the same concentration as in (b)?

3.11 **Silicon resistive sensors and their transfer functions.** A silicon resistive sensor has a nominal resistance of 2000 Ω at 25°C. To calculate its transfer function its resistance is measured at 0°C and 90°C and found to be 1600 Ω and 3200 Ω, respectively. Assuming the resistance is given by a second-order Callendar–Van Dusen equation, calculate the coefficients of the equation and plot the transfer function between 0°C and 100°C.

Thermistors

3.12 **Thermistor transfer function.** The transfer function of an NTC thermistor is best approximated using the Steinhart–Hart model in **Equation (3.14)** or **Equation (3.15)**. To evaluate the constants, a thermistor's resistance is measured experimentally at three temperatures, giving the following results: $R = 1.625$ kΩ at 0°C, $R = 938$ Ω at 25°C, and $R = 154$ Ω at 80°C.

a. Calculate the transfer function using the Steinhart–Hart model.

b. Using the resistance at 25°C as the reference temperature, calculate the transfer function using the simplified model in **Equation (3.12)**.

c. Plot the two transfer functions in the range 0°C–100°C and discuss the differences between them.

3.13 **Thermistor transfer function.** A new type of thermistor rated at 100 kΩ at 20°C is used to sense temperatures between −80°C and +100°C. It is expected that the transfer function is a second-order polynomial. To evaluate its transfer function, the resistance of the thermistor is measured at −60°C as 320 kΩ and at +80°C as 20 kΩ.

a. Find and plot the transfer function for the required span using a second-order polynomial.

b. Calculate the resistance expected at 0°C.

3.14 **Thermistor simplified transfer function.** The transfer function of a thermistor over the range 0°C–120°C is required. The thermistor is rated at 10 kΩ at 20°C. Three measurements are taken, at 0°C, 60°C, and 120°C, resulting in 24 kΩ, 2.2 kΩ, and 420 Ω, respectively. The simple exponential model in **Equation (3.12)** is used to derive the model. However, since the model only has one variable, β, one can choose any of the three temperatures to derive the transfer function.

a. Derive the transfer function using, in turn, the three measurements. Compare the values of β obtained.

b. Calculate the errors at the three points for the three transfer functions.

c. Plot the three transfer functions and compare them. Discuss the differences and the "proper" choice of a temperature in deriving the simplified model.

3.15 **Self-heat of a thermistor.** A thermistor with a nominal resistance of 15 kΩ at 25°C carries a current of 5 mA. At an ambient temperature of 30°C (measured with a thermocouple), the resistance of the thermistor is 12.5 kΩ. The current is now removed and the resistance of the thermistor drops to 12.35 kΩ. Calculate the error due to self-heat of the thermistor in degrees Celsius per milliwatt (°C/mW).

Thermoelectric sensors

3.16 Improper junction temperatures. A K-type thermocouple measures temperature T_1 and has reference T_r as shown in **Figure 3.32**. Calculate the reading of the voltmeter under the following conditions:

a. $T_1 = 100°C$, $T_r = 0°C$, and the junctions x–x' and y–y' are each in their own temperature zones. (c = chromel, a = alumel).

b. $T_1 = 100°C$, $T_r = 0°C$, and junctions y–y' are in a temperature zone. The junctions x–x' are not in a temperature zone with a temperature difference of 5°C (x is at the higher temperature).

c. Which reading is correct and what is the error in temperature using the incorrect reading?

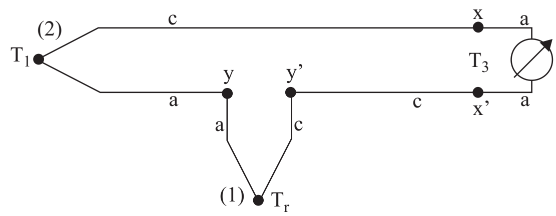


FIGURE 3.32 ■ Connection of a K-type thermocouple.

3.17 Extension of thermocouple wires. A K-type thermocouple measures temperature T_1 and has reference T_r as shown in **Figure 3.33**. Calculate the reading of the voltmeter under the following conditions:

a. $T_1 = 100°C$, $T_r = 0°C$, and the extension section is absent. The y–y' junctions are in a temperature zone. (c = chromel, a = alumel). The extension section is not present (i.e., x and w are the same point and x' and y' are the same point).

b. $T_1 = 100°C$, $T_r = 0°C$, and the extension section is present with the alumel wire on top and the chromel wire on the bottom. The junctions x–x', y–y', and z–z' are each in their own temperature zones, all three at 25°C. The junction w is at 20°C. Calculate the error in the reading of temperature T_1.

c. In an attempt to reduce the error, the extension is flipped so that now the chromel wire is on top and the alumel wire is on the bottom. Does that resolve the issue?

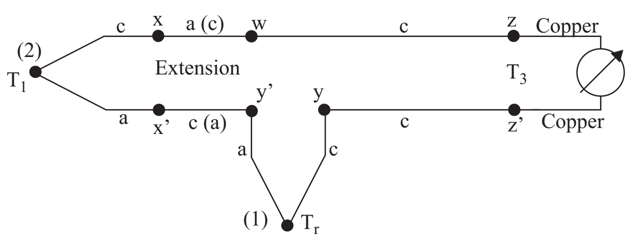


FIGURE 3.33 ■ Extension of thermocouple wires.

3.18 Reference junction with measured temperature. The configuration in **Figure 3.34** is used in a temperature-sensing system. The sensing and reference junctions are both K-type thermocouples. The sensing junction measures the temperature of molten glass at 950°C, whereas the reference thermocouple's temperature is measured as 54°C. The two connections marked as A and B are also within the same temperature zone as the reference junction. The temperature zone T_1 contains the connections to the measuring instrument.

a. Calculate the emf measured by the voltmeter using **Table 3.6**.

b. Show how that same emf can be obtained from the thermoelectric reference table for the K-type thermocouple by taking into account the emf of the reference junction and discuss the differences between the two methods.

c. Show that if the reference junction is held at 0°C the output is identical to that predicted by the thermoelectric reference table for temperature T.

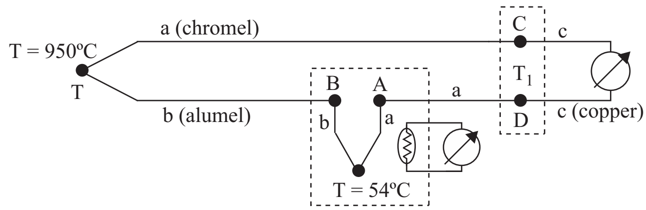


FIGURE 3.34 ■ Temperature sensing of the cold junction.

3.19 Thermoelectric reference emf and temperature for a type T thermocouple. The reference emf and reference temperature for copper-constantan (type T thermocouple) in the range 0°C–400°C are given as follows:

$$
\begin{aligned}
E = {} & 3.8748106364 \times 10^{1} T^{1} + 3.3292227880 \times 10^{-2} T^{2} + 2.0618243404 \\
& \times 10^{-4} T^{3} - 2.1882256846 \times 10^{-6} T^{4} + 1.0996880928 \times 10^{-8} T^{5} \\
& - 3.0815758772 \times 10^{-11} T^{6} + 4.5479135290 \times 10^{-14} T^{7} \\
& - 2.7512901673 \times 10^{-17} T^{8}\ \mu\text{V}
\end{aligned}
$$

$$
\begin{aligned}
T_{90} = {} & 2.5949192 \times 10^{-2} E^{1} - 7.602961 \times 10^{-7} E^{2} + 4.637791 \times 10^{-11} E^{3} \\
& - 2.165394 \times 10^{-15} E^{4} + 6.048144 \times 10^{-20} E^{5} - 7.293422 \times 10^{-25} E^{6}.
\end{aligned}
$$

a. Calculate the emf expected at 200°C using the first term (first-order approximation of the transfer function), first two terms (second-order approximation of the transfer function), first three terms, and so on until the complete eighth-order polynomial is used. Use a reference temperature of 0°C.

b. Calculate the error incurred in using reduced-order approximations compared to the exact value using all eight terms of the approximation. Plot the error as a function of the number of terms. What are your conclusions from these results.

c. Take the value found in (a) for the eighth-order approximation and calculate the reference temperature corresponding to the emf using one term in the expression for T (first-order approximation), first two terms (second-order approximation), first three terms, and so on up to six terms. Compare the results to the nominal temperature (200°C). What are the errors in the calculated temperatures with the various approximations? What are your conclusions from these calculations?

3.20 Cold junction compensation of an E-type thermocouple. Consider the cold junction compensation of a chromel-constantan thermocouple using a platinum RTD as shown in **Figure 3.11**. The RTD has a resistance of 120 Ω at 0°C and a TCR coefficient of 0.00385/°C. The relative Seebeck coefficient (sensitivity) for the E-type thermocouple at 0°C is 58.7 μV/°C (see **Table 3.3**).

a. Given a regulated voltage source of 5 V, calculate the resistance R_2 required for this type of thermocouple.

b. Calculate the error in temperature measurement at 45°C if the temperature zone is at $T = 25$°C. Use **Equation (3.20)** to calculate the emf for the T-type thermocouple.

c. Suppose the same configuration is used to sense a temperature of 400°C. What is the error? Discuss the sources of the error. Use the inverse polynomial for the T-type thermocouple in **Appendix B, Section B.4**, to calculate the emf due to $T_2 = 1{,}200$°C.

Semiconductor thermocouples

3.21 High-temperature thermopile. In remote areas where fuel, such as natural gas, is readily available, thermopiles are sometimes used to supply small amounts of power for specific needs, such as communication equipment and cathode protection of pipelines, among others. Consider the following example: A thermopile is needed to supply 12 V DC for emergency refrigeration using a thermopile. To do so, the hot junctions are heated to 450°C. The cold junctions are thermally connected to a conducting structure with fins, cooled by air, and expected to fluctuate between 80°C and 120°C depending on air temperature and wind speed. Because of the high temperatures involved, Peltier cells are not practical and it is proposed to use J-type thermocouples for the purpose.

a. Calculate the number of junctions needed to ensure a minimum output of 12 V.

b. What is the range of the output voltage?

3.22 Automotive thermogenerator. One of the more interesting attempts at using Peltier cells is in the replacement of alternators in trucks and, in the process, to recover some of the power loss through heat in the exhaust. The device is in the form of a cylindrical arrangement of junctions placed over the exhaust pipe, with the inner, hot junctions kept at the temperature of the exhaust pipe. The outer, cold junctions are kept at a temperature differential by a set of cooling fins on the outer surface of the cylindrical structure (or by circulating radiator coolant). Assuming that a minimal temperature difference of 60°C can be maintained between the hot and cold junctions, calculate the number of junctions needed to supply a minimum voltage of 27 V for trucks that operate at 24 V. The material used for construction of the junctions is carbon/silicon carbide because of its temperature range and high sensitivity (see **Table 3.9**).

Note: Prototypes capable of supplying about 5 kW have been built. However, there are some problems with this type of device. They are relatively expensive because of the need for high-temperature materials and can only supply power after the exhaust pipe has reached its normal operating temperature.

p-n junction temperature sensors

3.23 **Errors in *p-n* junction sensor.** There are two main errors that one should be aware of when using *p-n* junction sensors. One is the self-heat of the junction, the other is introduced by variations in the current through the diode. Consider a germanium diode with a known forward voltage drop of 0.35 V at a current of 5 mA and an ambient temperature of 25°C. The self-heat of the sensor is given in the device data sheet as 1.3 mW/°C.

a. Calculate the sensitivity of the sensor and the expected voltage reading at 50°C neglecting effects of self-heat.
b. Suppose that the current in the diode varies by ±10% due to variations in the power supply. Calculate the error in the measured temperature due to this variation as a percentage. Evaluate it at the values given above ($V_f = 0.35$ V, $T = 25$°C, $I = 5$ mA).
c. Calculate the error in the measured temperature due to self-heat at a junction current of 5 mA and an ambient temperature of 50°C.
d. Discuss these errors, their relative importance, and means of reducing them.

Optical and acoustical sensors

3.24 **Acoustic temperature sensing in seawater.** To sense the average water temperature close to the surface, an ultrasound transmitter and receiver are set at a distance 1 m apart and the time of flight of the ultrasound wave is measured using a microprocessor, as shown in **Figure 3.35**. The time Δt gives a direct indication of the temperature since the speed of sound in seawater is temperature dependent.

a. Calculate the sensitivity of the sensor.
b. Calculate and plot the measured time of flight as a function of temperature for the expected range of seawater temperatures between 0°C and 26°C.

Note: A sensor of this type is probably not something one would build, but if ultrasound measurements are being used for some other purpose the temperature can then be deduced as well.

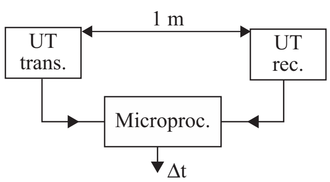


FIGURE 3.35 ■ Ultrasonic water temperature sensor.

3.25 Ultrasonic autofocusing and errors due to temperature variations. An ultrasonic sensor is used as a range meter for a camera to autofocus the lens. The method works based on measuring the time of flight of the ultrasound beam to the subject and back to the camera. Suppose the autofocus system is calibrated at 20°C.

a. Calculate the error in reading the distance as a percentage due to changes in the temperature of the air.

b. What is the actual distance measured at −20°C and at +45°C if the subject is 3 m away from the camera?

Thermomechanical sensors and actuators

3.26 The mercury thermometer. A mercury thermometer is made of glass as shown in **Figure 3.24**. The thermometer is intended for laboratory use with a scale of 0.5°C/mm and a range between 0°C and 120°C. If the thin tube is 0.2 mm in diameter, what is the volume of mercury necessary?

3.27 Gas temperature sensor. A gas temperature sensor is built in the form of a small container and a piston as in **Figure 3.22**. The piston is 3 mm in diameter. If the total volume of gas at 20°C is 1 cm^3, calculate the sensitivity of the sensors in millimeters per degree Celsius (mm/°C). Assume the pressure of the gas remains unchanged.

3.28 Fluid-filled Golay cell. A Golay cell is built as a cylindrical container with a flexible membrane, as shown in **Figure 3.36**. Its radius is $a = 30$ mm and its height is $h = 10$ mm. The membrane is stretched between the rim of the cylinder and a rigid disk of radius $b = 10$ mm. When the cell expands due to heating, the rigid disk lifts, stretching the membrane, and in so doing it creates a cone above the cylinder as shown in **Figure 3.36b**. The cell is filled with alcohol and the output is read optically as follows: A small mirror is attached to the membrane and a laser beam is reflected off the mirror. The cell in **Figure 3.36a** is shown at 0°C with the surface of the membrane perfectly flat. The reflected laser beam is read (using an optical sensor) on a scale at a distance of 60 mm from the mirror.

a. Calculate the sensitivity of the sensor for small changes in temperatures close to 0°C. *Note*: The reflection angle of the light equals the incidence angle, where the angles are measured with respect to the normal to the mirror at the

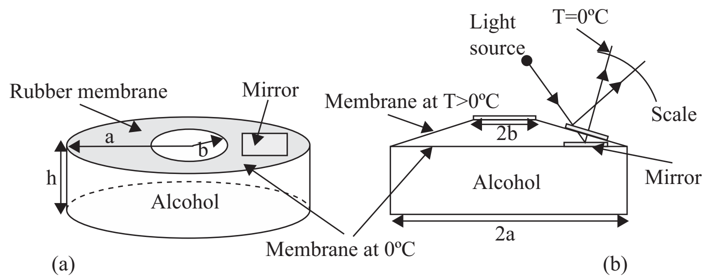


FIGURE 3.36 ■ A fluid-filled Golay cell. a. At 0°C. b. At T>0°C.

point of incidence. Since the sensed quantity is temperature (input) and the output is a linear length on the scale, sensitivity is given in millimeters per degree Celsius (mm/°C).

b. Assuming that the sensor is capable of distinguishing a beam separation of 0.1 mm on the scale, calculate the resolution of the Golay cell for small variations in temperature around 0°C.

3.29 **Piston-type Golay sensor.** A temperature sensor is made in the form of a glass container filled with air and a piston, as shown in **Figure 3.37**. The total volume of the gas at 0°C is 10 cm^3 and the piston's location indicates the temperature on a scale marked on the cylinder. Assume ideal gas behavior and no friction due to the piston (i.e., the internal pressure equals the external pressure).

a. If the external pressure is constant and equal to 1 atm (1013.25 mbar or 101,325 Pa), calculate the sensitivity of the sensor in degrees Celsius per millimeter displacement of the piston.

b. What is the maximum error in the reading if the external pressure changes from a low pressure of 950 mbar (95,000 Pa) to a high pressure of 1100 mbar (110,000 Pa) while the internal pressure stays at 1013.25 mbar (101,325 Pa).

c. What are the conclusions from (a) and (b)?

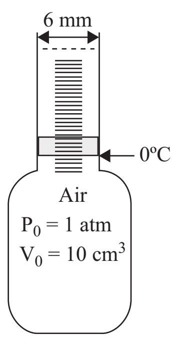


FIGURE 3.37 ■ Piston-type Golay sensor.

3.30 **Bimetal thermostat.** A thermostat is built as a bimetal bar and a snap switch as shown in **Figure 3.38** and used to control temperature in a small chamber. A snap switch operates by pressing against a strip spring, which when pushed beyond a certain point, snaps to a position that opens (or closes, depending on the type of switch) the contacts. When released, the contacts close (or open). The travel necessary is short, often less than 1 mm. In the case discussed here, the switch travel required is $d = 0.5$ mm. The bimetal bar is $t = 1$ mm thick made of iron (on the bottom) and copper (on the top).

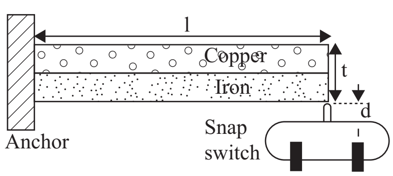


FIGURE 3.38 ■ Bimetal switch.

a. Assuming that at room temperature of 20°C the bar just touches the switch actuator and the thermostat must open the switch at 350°C, what is the minimum bar length needed.

b. If the distance *l* can be adjusted to a minimum of 25 mm, what is the highest temperature to which the thermostat can be set?

3.31 Coil bimetal thermometer. A coil bimetal thermometer is designed to operate in a range of 0°C–300°C. Assuming the inner diameter of the coil is 10 mm, calculate the length of the strip required to produce a 30° circular scale. The strip is made of copper (outer metal) and iron (inner metal). Assume that as the strip expands, all its coils maintain their diameter. The dial is moved directly by the inner coil.

3.32 The car thermostat: principle. A thermostat for use in a car engine is designed to open at 110°C. It does so through the use of a solid wax pellet in a cylinder that melts at a given temperature and as it does, expands considerably. The volume of the wax used expands by 5% as it melts. The configuration employed is a straight cylinder of diameter $d = 15$ mm with a piston connected directly to a disk that blocks the flow of water to the engine (**Figure 3.39**). The disk must move a distance $a = 6$ mm to fully open. Calculate the volume of the pellet necessary and the dimensions of the cylinder to accomplish this.

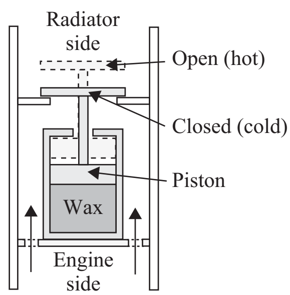


FIGURE 3.39 ■ The car thermostat: principle.

3.33 The car thermostat: a practical design. A practical configuration for a car thermostat is shown in **Figure 3.40**. Here the piston has been reduced in diameter to 3 mm, whereas the diameter of the cylinder remains the same as in **Problem 3.32**. In practical designs the cylinder moves down against a spring, whereas the piston itself is stationary. Calculate the volume necessary and the length of the actuating cylinder.

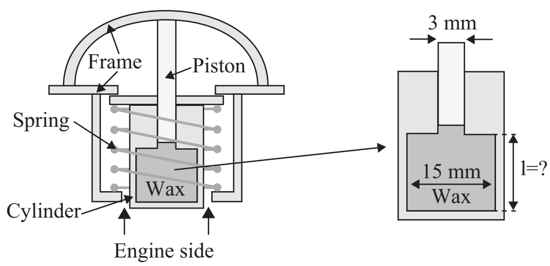


FIGURE 3.40 ■ A modified car thermostat.